

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 15:55:36 ; Search time 8.12473 Seconds
(without alignments)
3199.272 Million cell updates/sec

Title: US-10-723-552-3_COPY_585_2156

Perfect score: 2848
Sequence: 1 ATGGGACAGCGGTGACGAC.....CTCTAGAGAGATAAAGAT 1572

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10723552/runat 14022006 125146 13079/app.query.fasta_1
-DB=Issued Patents AA -QMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10723552 @CGN 1.1 161 @runat 14022006 125146 13079
-NCPU=6 -ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/iaa/5 COMB pep.*
- 2: /cgn2_6/ptodata/1/iaa/6 COMB pep.*
- 3: /cgn2_6/ptodata/1/iaa/H COMB pep.*
- 4: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE COMB pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2640	92.7	524	2	US-09-171-553B-4
2	1733	60.8	522	2	US-09-075-272-3
3	1733	60.8	622	2	US-09-075-272-2
4	1380.5	48.5	538	2	US-09-309-572-12
5	1380.5	48.5	538	2	US-09-718-096-12
6	1380.5	48.5	1737	2	US-09-309-572-13
7	1380.5	48.5	1737	2	US-09-718-096-13
8	1380	48.5	1312	2	US-09-554-572-26
9	1379.5	48.4	538	2	US-09-370-368-9
10	403.5	14.2	140	2	US-08-850-961-12
11	403.5	14.2	140	2	US-09-479-776-12
12	403.5	14.2	141	2	US-08-850-961-10

13	403.5	14.2	141	2	US-09-479-776-10
14	277	9.7	420	2	US-09-719-554-58
15	275.5	9.7	378	2	US-08-979-847B-122
16	275.5	9.7	398	2	US-08-979-847B-121
17	273	9.6	360	2	US-09-719-554-36
18	269	9.4	87	2	US-09-938-806A-7
19	267.5	9.4	352	2	US-08-979-847B-118
20	234	8.2	121	2	US-08-991-789A-2
21	234	8.2	121	2	US-09-062-451-2
22	234	8.2	121	2	US-09-598-326-2
23	234	8.2	121	2	US-09-289-198-2
24	234	8.2	121	2	US-09-429-755-2
25	234	8.2	121	2	US-09-699-295-2
26	234	8.2	121	2	US-09-534-825A-2
27	195	6.8	142	2	US-09-719-554-33
28	170.5	6.0	779	2	US-10-164-595-56
29	163.5	5.7	416	2	US-09-690-454-136
30	162.5	5.7	623	2	US-10-104-047-3378
31	161.5	5.7	1185	2	US-09-041-886-23
32	161.5	5.7	1185	2	US-09-538-092-1209
33	161	5.7	843	2	US-10-164-595-54
34	161	5.7	1184	2	US-09-266-225D-18
35	160.5	5.6	553	2	US-09-949-016-7961
36	159	5.6	1344	2	US-09-949-016-10925
37	157	5.5	1341	2	US-09-949-016-6890
38	154.5	5.4	420	2	US-09-902-540-13993
39	154	5.4	432	4	PCT-US95-04910-13
40	154	5.4	567	2	US-09-205-258-573
41	154	5.4	567	2	US-10-004-860-573
42	154	5.4	707	2	US-09-919-039-278
43	154	5.4	707	2	US-09-538-092-993
44	150	5.3	719	2	US-09-949-016-7766
45	150	5.3	919	2	US-09-949-016-6954

ALIGNMENTS

RESULT 1

US-09-171-553B-4

; Sequence 4, Application US/09171553B

; Patent No. 6756227

; GENERAL INFORMATION:

; APPLICANT: GALBRAITH, DANIEL N.

; APPLICANT: HAMORTH, CHRISTINE

; APPLICANT: LEES, GILLIAN M.

; APPLICANT: SMITH, KENNETH T.

; TITLE OF INVENTION: PORCINE RETROVIRUS

; FILE REFERENCE: CFV-5.01

; CURRENT APPLICATION NUMBER: US/09/171,553B

; CURRENT FILING DATE: 1999-02-08

; PRIOR APPLICATION NUMBER: PCT/GB97/01087

; PRIOR FILING DATE: 1997-04-18

; PRIOR APPLICATION NUMBER: GB 9702668.6

; PRIOR FILING DATE: 1997-02-10

; PRIOR APPLICATION NUMBER: GB 9608164.1

; PRIOR FILING DATE: 1996-04-19

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 524

; TYPE: PRT

; ORGANISM: Porcine retrovirus

US-09-171-553B-4

Alignment Scores:

Pred. No.: 3.64e-236

Score: 2640.00

Percent Similarity: 97.3%

Best Local Similarity: 94.3%

Query Match: 2

DB: 2

Length: 524

Matches: 495

Conservative: 16

Mismatches: 13

Indels: 2

Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x US-09-171-553B-4 (1-524)

QY 1 ATGGGACAGCGGTGACGACCCCTTTAGTTTGTACTCTCGACCATTTGGACTGAAGTTAAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20
QY 61 TCCAGGGCTCATAAATTTCTCAGTTTAAAGAGGACCTTGGCAGACCTTTCTGTGTC 120
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysGlyProTrpGlnThrPheCysAla 40
QY 121 TCTGAATGGCCGACATTCGATGTGGATGGCCATCAGAGGGGACCTTAAATTCGAGATT 180
Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
QY 181 ATCTTGGCTGTTAAAGCAGTTATTTTCAGACTGGACCCGCTCTCATCCCGATCAGGAG 240
Db 61 IleLeuAlaValLysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
QY 241 CCCTATATCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATGGTGTAAACCATGG 300
Db 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
QY 301 CTGAATAAGCCAGAACCCAGTCCCGAATTTCTGGCTCTTGGAGAGAAAAACAAACAC 360
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
QY 361 TCGGCTGAAAAGTCAAGCCCTC-TCCTCATATCTACCCGAGATTGAGGAGCCACCGGC 419
Db 121 SerAlaGluLysValGluProSerSerSerTyrLeuProArgProArgGlyAlaAlaAsp 140
QY 420 TTGGCCGGAAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGTCGCCGAGGGG 479
Db 141 LeuAlaGlyThrProThrCysSerProThrProLeuSerSerThrGlyCysGluGly 160
QY 480 ACCCTTTGCCCTCTCTGAGCTCCGGCGTGAGGGACCTCTGTCAGGAGCTCGAGCCG 539
Db 161 Thr-SerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerAr 180
QY 540 GAGGGGCCACCCCGGACGGACAGACGAGATCGGCATTTACCGTCGCCACGTAGCG 599
Db 180 GArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGl 200
QY 600 CCTCCACACCGGGGGCCAAATTCAGCCCTCCAGTATTTGGCCCTTTCTTCGCAGA 659
Db 200 YProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAs 220
QY 660 TCTCTAATTTGAAAACTAACCATTCCTCCCTTTCTCGAGAGATCCCAACGCTCACGGG 719
Db 220 pLeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGl 240
QY 720 GTTGTGGAGTCCCTATGTCTCTCACCGCTACTTTGGGATGATTCACACAGCTGCT 779
Db 240 YLeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLe 260
QY 780 GCAGACACTCTTCAACACGAGGACGAGAGAGAATTTCTTAGAGGCTAGAAAAATGT 839
Db 260 uGlnThrLeuPheThrThrGluLargGluArgIleLeuLeuGluAlaArgLysAsnVa 280
QY 840 TCCTGGGCCCGACGGCGCACCCGCGTTCAAAATGAGATTGAGTTGGGATTTCCCTT 899
Db 280 lProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe 300
QY 900 AACTCGCCCGTGGGACTACAAACCGCTGAAGGTAGGAGACCTTGAATCTATCG 959
Db 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrAr 320
QY 960 CCAGGCTCTGTGGGGGTCTCCGGGGCCCTCAAGACGGCCACTAATTTGGCTTAAGT 1019
Db 320 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVa 340
QY 1020 AAGAGAAGTGTACGGGACCGAATGAACCCCTCTGTGTTTCTTGAGAGGCTCTTGA 1079
Db 340 lArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGl 360

QY 1080 AGCTTTCAGGCGGTACACCCCTTTTGATCCCACTCAGAGGCCCAAAAGCCTCAGTGGC 1139
Db 360 uAlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAl 380
QY 1140 TTTGGCCTTTATAGGACAGTCAGCTTGGATATTAGAAAAGCTTCAGACACTGGAAAG 1199
Db 380 aLeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGl 400
QY 1200 GTTACAGGAGCTCAGTTACGTGATCTAGTGAAGGSCAGAGAAAGTATATTACAAAAG 1259
Db 400 YLeuGlnGluAlaGlnLeuArgAspLeuValArgGluAlaGlnLysValTyrTrpArgAr 420
QY 1260 GGACACACAAAGAAAGGGAACAAAGAAAAGAGAGAGAAAGAGAGGAAAGGAGGAAAG 1319
Db 420 gGluThrGluGluGlyGlnGlnArgLysGluLysGluLysGluArgGluGluArg 440
QY 1320 ACCTAATAACCGCAAGAGAAATTTGACTAAGATCTTGGCTGCTGCTGCTGCTGCTGCTG 1379
Db 440 gArgAspArgArgGlnGlnLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLy 460
QY 1380 AAGCAATACGAAAGAGAGAGAGATTTTAGGAAAATTAGGTACGCCCTCAGACAGTCAGG 1439
Db 460 sSerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGl 480
QY 1440 GAACCTGGCAATAGGACCCCACTCGACAGGACCAATGTGCATATTGTAAAAGAAAGAGG 1499
Db 480 YAsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGl 500
QY 1500 ACCTGGGCAAGAACTGCCCAAGAGGGGAAAACAAAGGACCAAGGATCTCTAGCTCTAGA 1559
Db 500 YHisTrpAlaArgAsnCysProLysGlyAsnLysGlyProLysValLeuAlaLeuGl 520
QY 1560 AGAAGATAAAGAT 1572
Db 520 uGluAspLysAsp 524
RESULT 2
US-09-075-272-3
; Sequence 3, Application US/09075272
; Patent No. 6136598
; GENERAL INFORMATION:
; APPLICANT: MILLER, A. DUSTY
; APPLICANT: WOLGAMOT, GREG
; APPLICANT: BONHAM, LYNN
; TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL
; TITLE OF INVENTION: PACKAGING CELL LINES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,272
; FILING DATE: 08-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,140
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: POOR, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: 14538A-003710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 522 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-09-075-272-3

Alignment Scores:

Pred. No.:	5,1e-152	Length:	522
Score:	1733.00	Matches:	340
Percent Similarity:	73.3%	Conservative:	61
Best Local Similarity:	62.2%	Mismatches:	96
Query Match:	60.8%	Indels:	50
DB:	2	Gaps:	12

US-10-723-552-3_copy_585_2156 (1-1572) x US-09-075-272-3 (1-522)

```

QY 1 ATGGGACAGCGGTGACGACCCCTCTAGTTTGACTCTCGACCAATTTGGACTGAAGTTAAA 60
Db 1 MetGlyGlnSerValSerThrProLeuSerLeuThrLeuGluHisTrpLysGluValys 20
QY 61 TCCAGGCGCTCATAATTTGTCAGTTTCAGTTTAAGAGGACCTTGCAGACTTCTGTGTC 120
Db 21 IleArgAlaHisAsnGlnSerValGluValArgLysGlyProTrpGlnThrPheCysAla 40
QY 121 TCTGAATGGCGGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTTCTGAGATT 180
Db 41 SerGluTrpProThrPheGlyValGlyTrpProGluGlyAlaPheAspLeuSerLeu 60
QY 181 ATCTGCGCTGTTAAAGCAGTATTATTTTCAGACTCGGCGCTCTCATCCGATCAGAG 240
Db 61 IleAlaAlaValArgArgIleValPheGlnGlu---GluGlyGlyHisProAspGlnIle 79
QY 241 CCCTATATCTTACGTGCGAAGATTTCGAGAGGATCTCCGCGCATCGSETTAACCATGG 300
Db 80 ProTyrIleValThrTrpGlnSerLeuValGlnPheProSerProSerTrpValLysProTrp 99
QY 301 CTGAATAAGCCACAGAAAGCCAGGTCCCGCAATTTCTGGCTCTTGGAGAGAAAAACAAAC 360
Db 100 -----ThrProAsnProSerLysLeuThrValAlaValAlaGlnSerAsp 114
QY 361 TCGGCTGAAAAGTCAAGCCCTCT-CTCATATCTACCCCGAGATTGAG-----408
Db 115 AlaAlaGluLysSerGlyProSerAlaProProLysIleTyrProGluIleAspLeu 134
QY 409 -----GAGCCACCGCTTGGCGGAACCCCAATCTGTTCC-----CCA 447
Db 135 LeuTrpMetAspSerGlnProProTyrProLeuProGlnGlnProProAlaAlaAla 154
QY 448 CCCCCTTATCTGGCACAGGGTGGCGGAGGGGACCTTTGCCCCCTCTCGAGCTCCGCGG 507
Db 155 ProProValAlaAla-----ProGlnProGluProThrAlaSerGly 168
QY 508 GTGAGGGACCTGCTCGAGGACTCGAGCGGAGGGGCGCCACCCCGAGCGG-----561
Db 169 AlaGlnGlyProAlaGlyGlyThrArgSerArgArgGlyArgSerProAlaGluGluGly 188
QY 562 ---ACACACGAGATCGCGACATTACCGCTCGCAGCGTAC---GGCCCTCCACACCGGG 615
Db 189 GlyProAspSerThrValAlaLeuProLeuArgAlaHisValGlyGlyProThrProGly 208
QY 616 ---GGCCAAATTGACGCCCTCCAGTATTGGCCCTTTCTCTGAGATCTCTATAATTGG 672
Db 209 ProAsnAspLeuLeuProLeuGlnTyrTrpProPheSerSerSerAspLeuTyrAsnTrp 228
QY 673 AAAACTAACCATCCCTTTCTCGAGGATCCCAACGCTCAAGGGGTTGGTGGAGTCC 732
Db 229 LysThrAsnHisProProPheSerGluAsnProSerGlyLeuThrGlyLeuLeuGluSer 248
QY 733 CTTATGTTCTTCCACGACCTACTTGGGATGATTGTCAACAGCTGTGCGACACTCTTC 792
Db 249 LeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnValLeuPhe 268

```

```

QY 793 ACAACGAGGAGCGAGAGAAATTTCTATTAGAGCTAGAAAAAATGTTCTCTGGGCGCCGAC 852
Db 269 ThrThrGluGluArgGluArgIleLeuMetGluAlaArgLysAsnValLeuGlyGluAsp 288
QY 853 GGGCGACCCACGCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTAACTCGCCCCCGGT 912
Db 289 GlyThrProThrAlaLeuProAsnLeuValAspGluAlaPheProLeuAsnArgProAsn 308
QY 913 TGGGACTACAAACACGCGTAGAGGTAGGAGAGCTTGAAAAATCTATCGCAGGCTCTGGTG 972
Db 309 TrpAspTyrAsnThrAlaGluGlyArgGlyArgLeuValTyrArgArgThrLeuVal 328
QY 973 GCGGGTCTCCGGGCGGCTCAAGACGCGCCCACTAATTTGGCTAAGGTAAAGAGAAGTGATG 1032
Db 329 AlaGlyLeuArgGlyAlaAlaArgProThrAsnLeuAlaValArgGluValLeu 348
QY 1033 CAGGGACCGAATGAACCCCTCTGTTTCTTTCAGAGGCTCTTGGAAAGCTTTCAGGCGG 1092
Db 349 GlnGlyGlnThrGluProProSerValPheLeuGluArgLeuMetGluAlaTyrArgArg 368
QY 1093 TACACCCCTTTTGATCCACCTCAGAGCGCCCAAAAGCTCAGTGGCTTTGCGCTTTTATA 1152
Db 369 TyrThrProPheAspProSerSerGlyGlnLysAlaValAlaMetAlaPheIle 388
QY 1153 GGACAGTCAGCCTTTGGATATTAGAAAGAGCTTCAGAGACTTGAAGGGTTACAGAGGCT 1212
Db 389 GlyGlnSerAlaProAspIleLysLysLeuGlnArgLeuGluGlyLeuGlnAspTyr 408
QY 1213 GAGTTAGCTGATCTAGTGAAGGAGGAGAGAGAAAGTATATTACAAAAGGAGACAGAGAA 1272
Db 409 ThrLeuGlnAspLeuValLysGluAlaGluLysValTyrHisLysArgGluThrGluGlu 428
QY 1273 GAAAGGAAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1332
Db 429 GluArgGlnGluArgGluLysGluValGluGluArgGluAsnArgArgArgArg 448
QY 1333 CAAGAGAGAAATTTGACTTAAGATCTTGGCTGCGAGTGTGTAAGGAGAAAGCAATACGGA 1392
Db 449 GlnGluArgAsnLeuSerLysIleLeuAlaValIleAsnAspArg-----464
QY 1393 AGAGAGAGATTTTAGGAAAAATTAGTCTAGGCGCTTAGACAGTCAAGGAACTGGGCAAT 1452
Db 465 -----GlnSerGluLysGlyArgThrGlyPheLeuGlyAsn 476
QY 1453 AGG-----ACCCCACTCGACAGGACCATGTGCATAT 1485
Db 477 ArgAlaValLysProProGlyGlyArgLysThrProLeuGluLysAspGlnCysAlaPhe 496
QY 1486 TGTAAGAGAGAGAGACACTGGGCAAGAACTGCCCAAGAGGAGGAGAAACAAAGGACCAAGG 1545
Db 497 CysLysGluLysGlyHisTrpAlaLysAspCysProLysLys---ArgArgGlnPheLys 515
QY 1546 ATCCTAGCTCTAGAAGAGAT 1566
Db 516 ValLeuThrLeuGluAspAsp 522

```

RESULT 3

US-09-075-272-2

; Sequence 2, Application US/09075272

; Patent No. 6136598

; GENERAL INFORMATION:

; APPLICANT: MILLER, A. DUSTY

; APPLICANT: WOLGAMOT, GREG

; APPLICANT: BOHAM, LYNN

; TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL

; TITLE OF INVENTION: PACKAGING CELL LINES

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

```

; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,272
; FILING DATE: 08-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,140
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: 14538A-003710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-075-272-2

Alignment Scores:
Pred. No.: 5,51e-152 Length: 622
Score: 1733.00 Matches: 340
Percent Similarity: 73.3% Conservative: 61
Best Local Similarity: 62.2% Mismatches: 96
Query Match: 60.8% Indels: 50
DB: 2 Gaps: 12

US-10-723-552-3_COPY_585_2156 (1-1572) x US-09-075-272-2 (1-622)
QY 1 ATGGACAGACGGTGACACCCCTCTTAGTTGACTCTCGACCAATGACCTGAAGTTAAA 60
DB 101 MetGlyGlnSerValThrProLeuSerLeuThrLeuGluHisTrpLysGluValLys 120
QY 61 TCCAGGGCTCATATTTGTCAGTTCAGTTAAGAGGACCTTGCCAGACCTTCTGTC 120
DB 121 IleArgAlaHisAsnGlnSerValGluValArgLysGlyProTrpGlnThrPheCysAla 140
QY 121 TCTGAATGCGACATTCGATGTCGATGGCCATCAGAGGGGACCTTTTAATTCAGATT 180
DB 141 SerGluTrpProThrPheGlyValGlyTrpProGluGlyAlaPheAsnLeuSerLeu 160
QY 181 ATCTGGCTGTATAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
DB 161 IleAlaAlaValArgArgIleValPheGlnGlu---GluGlyGlyHisProAspGlnIle 179
QY 241 CCTATATCTTACCTGTCAGATTTGGCAGAGGATCTCCGCCATGGTAAACCATGG 300
DB 180 ProTyrIleValThrTrpGlnSerLeuValGlnPheProSerTrpValLysProTrp 199
QY 301 CTGAATAAGCCAAGAAGACCGGTCCTCCGAAATTCCTGGCTCTGGAGAGAAAACAAACAC 360
DB 200 -----ThrProAsnProSerLysLeuThrValAlaValAlaGlnSerAsp 214
QY 361 TCGGCTGAAAAGTCAAGCCCTCT-----CTCATATCTACCCCGAGATTGAG----- 408
DB 215 AlaAlaGluLysSerGlyProSerAlaProProLysIleTyrProGluIleAspAspLeu 234
QY 409 -----GAGCCACGGCTGGCCGGGAACCCCATCTGTTCCC-----CCA 447
DB 235 LeuTrpMetAspSerGlnProProTyrProLeuProGlnGlnProProAlaAlaAla 254
QY 448 CCCCTTATCTGGCACAGGGTCCCGAGGGGACCTTTGGCCCTCTCTGGAGCTCCGGCG 507
DB 1

```

```

255 ProProValAlaAla-----ProGlnProGluProThrAlaSerGly 268
QY 508 GTGAGGACCTGCTGCAGGACTCGAGCGGGAGGGCGCCACCCCGAGCGG----- 561
DB 269 AlaGlnGlyProAlaGlyGlyThrArgSerArgArgGlyArgSerProAlaGluGly 288
QY 562 ---ACAGACGAGATCGCGACATTACCGCTCGCGCACGTAC---GGCCCTCCACACCGGG 615
DB 289 GlyProAspSerThrValAlaLeuProLeuArgAlaHisValGlyGlyProThrProGly 308
QY 616 ---GGCCAATTGACAGCCCTCCAGTATTGGCCCTTTCTCTCGAGATCTCTATAATTGG 672
DB 309 ProAsnAspLeuIleProLeuGlnTyrTrpProPheSerSerAspLeuTyrAsnTrp 328
QY 673 AAAACTAACCATCCCTTTCTCGAGGATCCCAACGCCCTCAGCGGTTTGGTGGAGTCC 732
DB 329 LysThrAsnHisProProPheSerGluAsnProSerGlyLeuThrGlyLeuGluSer 348
QY 733 CTTATGTTCTCTCACAGCCTACTTGGGATGTTCAACAGCTGCTGCAGACACTCTTC 792
DB 349 LeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnValLeuPhe 368
QY 793 ACAACCGAGGAGCAGAGAGAAATTTCTATTAGAGCTAGAAAAAATGTTCTCGGGCCGAC 852
DB 369 ThrThrGluGluArgGluArgIleLeuMetGluAlaArgLysAsnValLeuGlyGluAsp 388
QY 853 GGGGACCCACCGCGTTCAAAATCAGATTGACATGGGATTTCCCTTAACCTCGCCCGGT 912
DB 389 GlyThrProThrAlaLeuProAsnLeuValAspGluAlaPheProLeuAsnArgProAsn 408
QY 913 TGGGACTACAACACCGCTGAAGGTAGGAGAGCTTTGAAAAATCTATCGCAGGCTCTGCTG 972
DB 409 TrpAspTyrAsnThrAlaGluGlyArgGlyArgLeuValTyrArgThrLeuVal 428
QY 973 GCGGCTCTCGGGCGGCTCAAGACGCGCCACTAATTTGGTAAAGTAAAGAGAGTANG 1032
DB 429 AlaGlyLeuArgGlyAlaAlaArgArgProThrAsnLeuAlaLysValArgGluValLeu 448
QY 1033 CAGGACCGAATGAACCCCTCTGTTTCTTGAGAGGCTCTTGGAAGCTTCAGGCGG 1092
DB 449 GlnGlyGlnThrGluProProSerValPheLeuGluArgLeuMetGluAlaTyrArg 468
QY 1093 TACACCCCTTTGATCCACCTCAGAGCCCAAAAGCCTCAGTGGCTTTGGCCCTTTATA 1152
DB 469 TyrThrProPheAspProSerSerGlyGlyGlnLysAlaAlaValAlaMetAlaPheIle 488
QY 1153 GCAGCTCAGGCTTTGGATATTAGAAAGAGCTTCAGACACTCGAAGGTTACAGAGGCT 1212
DB 489 GlyGlnSerAlaProAspIleLysLysLysLeuGlnArgLeuGlyLeuGlnAspTyr 508
QY 1213 GAGTTACGTGATCTAGTGAAGGACAGAGAAAGTATATTACAAAGGGAGACAGAGAA 1272
DB 509 ThrLeuGlnAspLeuValLysGluAlaGlyLysValTyrHisLysArgGlnThrGluGlu 528
QY 1273 GAAAGGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1332
DB 529 GluArgGlnGluArgGlyLysGlyValGluGluArgGluAsnArgAspArgArg 548
QY 1333 CAAGAGAGAAATTTGACTTAAGATCTTGGCTCGAGTGGTGAAGGAGAAAGCAATACGAA 1392
DB 549 GlnGluArgAsnLeuSerLysIleLeuAlaValIleAsnAspArg----- 564
QY 1393 AGAGAGAGATTTTAGGAAATTAGTCAGGCCCTCAGCCCTAGACAGTCAAGGAACTGGGCAAT 1452
DB 565 -----GlnSerGluLysGlyArgThrGlyPheLeuGlyAsn 576
QY 1453 AGG-----ACCCCACTCGACAGAGACCAATGTGCATAT 1485
DB 577 ArgAlaValLysProProGlyGlyArgLysThrProLeuGluLysAspGlnCysAlaPhe 596
QY 1486 TGTAAGAGAGAGGACACTGGGCAAGGAATCTGCCCAAGAGAGGAGAAACAAAGGACCAAG 1545
DB 597 CysLysGluLysGlyHisTrpAlaLysAspCysProLysLys---ArgArgGlnPheLys 615

```

```

Qy 1546 ATCTAGCTCTAGAGAAGAT 1566
Db :||| |||||:|
616 ValLeuThrLeuGluAspAsp 622

RESULT 4
US-09-309-572-12
; Sequence 12, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag protein
US-09-309-572-12

Alignment Scores:
Pred. No.: 2,61e-119 Length: 538
Score: 1380.50 Matches: 283
Percent Similarity: 62.2% Conservative: 69
Best Local Similarity: 50.0% Mismatches: 141
Query Match: 48.5% Indels: 73
DB: 2 Gaps: 13

US-10-723-552-3_COPY_585_2156 (1-1572) x US-09-309-572-12 (1-538)
Qy 1 ATGGGACAGCGGTGACGACCCCTCTTAGTTGACTCTCGACCAATTGGAGCTGAAGTTAAA 60
Db :||| |||||:|
1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20

Qy 61 TCCAGGCTCATATTTGTCAGTTTCAGTTTAAGAGGACCTTGGCAGACTTCTGTGTC 120
Db :||| |||||:|
21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgTrpValThrPheCysSer 40

Qy 121 TCTGAATGGCGGACATTCGATTTGGATGGCATCAGAGGCGACTTTAATCTGAGATT 180
Db :||| |||||:|
41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60

Qy 181 ATCTGGCTGTAAAGCAGGTATTTTTCAGACTCGACCCGCTCTCATCCCGATCAGGAG 240
Db :||| |||||:|
61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80

Qy 241 CCTATATCTTACGTGGCAGATTTCGCGAGAGATCTCCGCGCATGGTTAAACCATGG 300
Db :||| |||||:|
81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProTyrValLysProPhe 100

Qy 301 CTGAATAAGCAAGAACCGAGTCCCGGATTCGTGCTCTTGAGAGAAAAACAACAC 360
Db :||| |||||:|
101 ValHis---ProLysProProProProLeuProProSerAlaProSerLeuProLeuGlu 119

Qy 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTTACCCCGAGATTTAGGAGCCCA---CCG 417
Db :||| |||||:|
120 ProProArgSerThrProProArgSerSerLeuTyrProAlaLeuThrProSerLeuGly 139

Qy 418 GTTGGCGGAAACCCCAA-----TCT 438
Db :||| |||||:|
140 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuIleAspLeuLeuThr 159

Qy 439 GTTCCCGCCCTTATCTGGCACAGGTTGCCGAGGGGACCTTTGGCCCTCTCT--- 495
Db :||| |||||:|
160 GluAspProProTyr-----ArgAspProArgProProProSer 173

Qy 496 -----GGAGCTCCCGCGGTGGAGGACCT----- 519
Db :||| |||||:|
174 AspArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyGluAlaProAspProSerPro 193

Qy 520 ---CCTCGAGGACTCGAGCCGAGGGCGCCACCCCGAGCGGACGAGACAGATCCGC 576
Db :||| |||||:|
194 MetAlaSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrThrSerGln 213

Qy 577 ACATTACCGCTGGCGACGTACGGCCCTCCACACCGGGGGCCAAATTCGACGCCCTCCAG 636
Db :||| |||||:|
214 AlaPheProLeuArgAlaGlyGly-----AsnGlyGln-----LeuGln 226

Qy 637 TATTGGCCCTTTCTCTCGAGATCTCTATAATTGAAAACTAACCACTCCCTTTCTCG 696
Db :||| |||||:|
227 TyrTrpProPheSerSerAspLeuTyrAsnTrpLysAsnAsnProSerPheSer 246

Qy 697 GAGGATCCCAACGCTCAGGGGTGGTGGAGTCCCTTATGTCTCTCACCAGCTACT 756
Db :||| |||||:|
247 GluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuIleThrHisGlnProThr 266

Qy 757 TGGGATGATTGTCAACAGCTGTGCACACACTCTTCAACCGAGGAGCGAGAGAGATT 816
Db :||| |||||:|
267 TrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgVal 286

Qy 817 CTATTAGAGCTAGAAAAAATGTTCTTGGGCGCGACCGGCGACCCAGCGGTTCAGAAAT 876
Db :||| |||||:|
287 LeuLeuGluAlaArgLysAlaValArgGlyAspAspGlyArgProThrGlnLeuProAsn 306

Qy 877 GAGATTGCATGGGATTTCCCTTAACCTCGCCCGGTTGGGACTCAACACCGCTGAAGGT 936
Db :||| |||||:|
307 GluValAspAlaAlaPheProLeuGluArgProAspTrpAspTyrThrThrGlnAlaGly 326

Qy 937 AGGAGAGCTTTGAAAAATCTATCGCCAGCTCTGTGTGGCGGTCTCCGGGGCGCTCAAGA 996
Db :||| |||||:|
327 ArgAsnHisLeuValHisTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 346

Qy 997 CGGCCCACTAATTTGGCTAAGTAAGAGAAGTGAATGACGAGGACCGAATGAACCCCTCT 1056
Db :||| |||||:|
347 SerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer 366

Qy 1057 GTTTTCTTTCAGAGGCTTGGAGCCTTCAGGCGGTACACCCCTTTTGATCCACCTCA 1116
Db :||| |||||:|
367 AlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp 386

Qy 1117 GAGGCCCAAAAGCCTCAGTGGCTTTGGCCCTTTATAGGACAGTCAGCTTGGATATTAGA 1176
Db :||| |||||:|
387 ProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly 406

Qy 1177 AAGAAGCTTCAGAGACTGGAAGGTTTACAGGAGGCTGAGTTAGCTAGTGTGAAGGAG 1236
Db :||| |||||:|
407 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 426

Qy 1237 GCAGAGAACTATTTACAAAAGGAGACAGAGAAGAAAGGGAACAAAGAAAGAGAGA 1296
Db :||| |||||:|
427 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgIleArgArg 446

Qy 1297 GAAAGAGAGAAAGGAGGAAGA-----CGT 1323
Db :||| |||||:|
447 GluThrGluGluLysGluGluArgArgThrGluAspGluGlnLysGluLysGluArg 466

Qy 1324 AATAAAGCGCAAGAGAAATTTGACTTAAGATCTTGGCTCAGTGGTTGAAGGGAAGG 1383
Db :||| |||||:|
467 AspArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys 486

Qy 1384 AATACGGAAGAGAGAGAGATTTTAGGAAATTTAGGTTCAGGCCCTAGACAGTCAGGGAAC 1443
Db :||| |||||:|
487 Gln-----AspArgGlnGlyGlu 493

Qy 1444 CTGGGCAATAGGACCCCTCGACAAAGCAATGTGCAATTGTAATGTAAGAAAGAGGACAC 1503
Db :||| |||||:|
494 ---ArgArgArgSerGlnLeuAspArgAspGlnCysAlaTyrCysLysGluLysGlyHis 512

Qy 1504 TGGGCAAGGAACCTGCCCAAGAG-----GGAAACAAAGGACCAAGG----- 1545

```

```

Db      513 TrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArgGlyProArgGlyProArgGlyProGlnThrSer 532
Qy      1546 ATCTAGCTCTAGAGAA 1563
Db      533 LeuLeuThrLeuAspAsp 538

RESULT 5
US-09-718-096-12
; Sequence 12, Application US/09718096
; Patent No. 6589763
; GENERAL INFORMATION:
; APPLICANT: Von Laer, Meike-Dorothee
; TITLE OF INVENTION: RETROVIRAL HYBRID VECTORS PSEUDOTYPED WITH LCMV
; FILE REFERENCE: 35-195
; CURRENT APPLICATION NUMBER: US/09/718,096
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: DE 19856463.5
; PRIOR FILING DATE: 1998-11-26
; PRIOR APPLICATION NUMBER: EP 99250415.9
; PRIOR FILING DATE: 1999-11-25
; PRIOR APPLICATION NUMBER: US 09/309,572
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag protein
US-09-718-096-12

Alignment Scores:
Pred. No.: 2,61e-119 Length: 538
Score: 1380.50 Matches: 283
Percent Similarity: 62.2% Conservative: 69
Best Local Similarity: 50.0% Mismatches: 141
Query Match: 48.5% Indels: 73
DB: 2 Gaps: 13

US-10-723-552-3_COPY_585_2156 (1-1572) x US-09-718-096-12 (1-538)
Qy      1 ATGGGACAGCGGTGACGACCCCTTCTAGTTGACTCTCGACCATGAGCTGAAGTAAA 60
Db      1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
Qy      61 TCCAGGGCTCATAATTGCTCAGTTTCAGGTTAAGAGGACCTTGGCAGACTTCTGTGTC 120
Db      21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgGTrpValThrPheCysSer 40
Qy      121 TCTGAATGCCGACATTCGATGTTGGATGGCCATCAGAGGGAGCTTTAAATTCGAGATT 180
Db      41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
Qy      181 ATCTGGCTGTAAAGCAGTATTTTTCAGACTGGACCGGCTCTCATCCCGATCAGAG 240
Db      61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
Qy      241 CCTATATCTTACGTGGCAGATTGTGGCAGAGGATCTCCGCCATGGTGGTTAAACCATGG 300
Db      81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProProTrpValLysProPhe 100
Qy      301 CTGAATAAGCCAAGAACCGAGTCCCGAATTCCTGGCTCTTGGAGAGAAACAAACAC 360
Db      101 ValHis---ProLysProProProLeuProProProSerAlaProSerLeuProLeuGlu 119
Qy      361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATGAGGACCA---CCG 417
Db      120 ProProArgSerThrProProArgSerSerLeuTyrProAlaLeuThrProSerLeuGly 139
Qy      418 GCTTGGCCGGAACCCCAA---TCT 438

```

```

Db      140 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuLysLeuLeuLeuThr 159
Qy      439 GTTCCCCCACCCTTATCTGGCACAGGGTCCCGAGGGGACCTTTTGCCTCCT--- 495
Db      160 GluAspProProProTyr-----ArgAspProA-gProProProSer 173
Qy      496 -----GGAGCTCCGGCGGTGGAGGACCT----- 519
Db      174 AspArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyGluAlaProAspProSerPro 193
Qy      520 ---GCTCAGGAGACTCGAGCCGAGGGCGCCACCCCGAGCGGACAGACGAGATCGCG 576
Db      194 MetAlaSerArgLeuArgGlyArgGluProProValAlaAspSerThrThrSerGln 213
Qy      577 ACATTACCGTGGCCACGTACGGCCCTCCACACCGGGGGCCCAATTGCAGCCCTCCAG 636
Db      214 AlaPheProLeuArgAlaGlyGly-----AsnGlyGln-----LeuGln 226
Qy      637 TATTGGCCCTTTTCTTCTGCAGATCTCTATAATTGGAAAACTAACCATCCCTTTCTCG 696
Db      227 TyrTrpProPheSerSerSerAspLeuTyrAsnTrpLysAsnAsnProSerPheSer 246
Qy      697 GAGGATCCCAACGCCCTCAGCGGGTGGTGAGTCCCTTANGTCTCTCACACGCTACT 756
Db      247 GluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuLeuThrHisGlnProThr 266
Qy      757 TGGGATGATTGTCAACAGCTCTCCAGACACTCTTCACACCGAGGAGCGAGAGAGATT 816
Db      267 TrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgVal 286
Qy      817 CTATTAGAGGCTAGAAAAATGTTCTCTGGGCGCCAGCGGCGACCCAGCGGTTTCAAAAT 876
Db      287 LeuLeuGluAlaArgLysAlaValArgGlyAspAspGlyArgProThrGlnLeuProAsn 306
Qy      877 GAGATTGACATGGGATTTCCTTAACTCGCCCCGGTTGGGACTACAAACACGGCTGAAGT 936
Db      307 GluValAspAlaAlaPheProLeuGluArgProAspTrpAspTyrThrThrGlnAlaGly 326
Qy      937 AGGAGAGCTTGAATAATCTATCGCAGGCTCTGTGGGGGTCTCCGGGGCGCTCAGA 996
Db      327 ArgAsnHisLeuValHisTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 346
Qy      997 CGGCCCACTAATTGGCTAAGTAAGAGAACTGATCGAGGAGCAGCAATGAACCCCTCT 1056
Db      347 SerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer 366
Qy      1057 GTTTTCTTGAGAGCTCTTGAAGCCTTCAGCGGTACACCCCTTTTGATCCCACTCA 1116
Db      367 AlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp 386
Qy      1117 GAGGCCCAAAAGCCTCAGTGGCTTTGGCCTTTATAGACAGTGCAGCTTGGATTATAGA 1176
Db      387 ProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly 406
Qy      1177 AAGAAGCTTCAGACACTGGAAGGGTTCAGAGGCTGAGTTACGTGACTAGTGAAGGAG 1236
Db      407 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 426
Qy      1237 GCAGAGAAGTATATTACAAAAGGAGACAGAGAAGAAAGGAAACAAAGAAAGAGAGA 1296
Db      427 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgIleArgArg 446
Qy      1297 GAAAGAGAGGAAAGGAGGAAGA-----CGT 1323
Db      447 GluThrGluGluLysGluGluArgArgThrGluAspGluGlnLysGluLysGluArg 466
Qy      1324 AATAAACCGCAAGAGAAGATTTGACTAAGATCTTGGCTGCGAGTGGTTGAAGGAAAGC 1383
Db      467 AspArgArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys 486
Qy      1384 AATACGGAAGAGAGAGAGATTATAGGAAATTTAGGTACAGCCCTTAGACAGTCAGGGAAC 1443
Db      487 Gln-----AspArgGlnGlyGlyGlu 493

```

QY	1444	CTGGC	CAATAGG	ACCC	CCTCG	ACAAG	CAAC	AAATGT	GCATAT	TCTTA	AGAAAG	AGG	CAC	1503	
DB	494	---	Arg	Arg	Ser	Gln	Leu	Asp	Arg	Asp	Gln	Cys	ala	512	
			---	Arg	Arg	Ser	Gln	Leu	Asp	Arg	Asp	Gln	Cys	ala	512
			---	Arg	Arg	Ser	Gln	Leu	Asp	Arg	Asp	Gln	Cys	ala	512
QY	1504	TGGC	AAGG	AACTG	CCCC	AGA	AG-	-----	GGAA	CAAA	AGAC	CA	AGG-	1545	
DB	513	Trp	Ala	lys	Asp	Cys	Pro	Lys	Pro	Arg	Gly	Pro	Arg	532	
QY	1546	ATC	TAG	CTCT	AG	AGAA	1563								
DB	533	Leu	Leu	Thr	Leu	Asp	Asp	538							

RESULT 6

```

US-09-309-572-13
; Sequence 13, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patencin ver. 2.0
; SEQ ID NO 13
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag-pol protein
US-09-309-572-13

```

Alignment Scores:	4.36e-119	Length:	1737
Pred. No.:	1380.50	Matches:	283
Score:	62.2%	Conservative:	69
Percent Similarity:	50.0%	Mismatches:	141
Best Local Similarity:	48.5%	Indels:	73
Query Match:	2	Gaps:	13
DB:			

US-10-723-552-3 COPY 585 2156 (1-1572) x US-09-309-572-13 (1-1737)

1	ATGGGACAGACGGTCAGCACCCCTCTTAGTTTGACTCTCGACCAATTGGACTGAAGTTAAA	60
1	MetGlyGlnThrValThrThrProLeuSerLeuThrLeuGlyHisTrpLeuAspValGlu	20
61	TCCAGGGCTCATAAATTCTGACTGAGTTAAGAGGGACCTTGGCAGACTTTCTCTGTC	120
21	ArgIleAlaHisAsnGlnSerValAspValLysLysArgArgTrpValThrPheCysSer	40
121	TCTGAATGGCCGACATTTCGATTTGGATGGCCATCAGAGGGGACCTTTAATCTCGAATT	180
41	AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu	60
181	ATCTCGGCTGTAAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCGATCAGGAG	240
61	IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal	80
241	CCCTATATCTTACCTGTGCAAGATTGCGAGAGATCTCCGCCATGGGTAAACCATCG	300
81	ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProTrpValLysProPhe	100
301	CTGAATAAGCCCAAGAACCGAGTCCCAGATTCTGGCTCTGGAGAGAAAAACAAAC	360
101	ValHis---ProLysProProProLeuProProSerAlaProSerLeuProLeuGlu	119
361	TGGGCTGAAAAGTCAGCCCTCTCTCATATCTACCCCGAGATTGGAGGCA---CCG	417
120	ProProArgSerThrProProArgSerSerLeuThrProAlaLeuThrProSerLeuGly	139

QY	418	GC	TTGGCGCGGAACCCCAA	----	-----TCT	438
			:			
Db	140	Ala	LyProLyProGlnValLeuSerAspSerGlyGlyProLeuIleAspLeuLeuThr		---	159
QY	439	GT	TCCCCACACCCCTTATCTGSCACAGGTGCGCGAGGGAGCCCTTTGCCCTCT	---	---	495
Db	160	Glu	AspProProProTyr	-----	ArgAspProArgProProProSer	173
QY	496	-----	-----GGAGCTCCGCGGTGGAGGGACCT	-----	-----	519
Db	174	Asp	ArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyGluAlaProAspProSerPro		193	
QY	520	---	GCTGAGGAGCTCGAGCGGAGGGGCGCCACCCCGGAGCGGACAGACAGATCGG		576	
Db	194	Met	AlaSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrThrSerGln		213	
QY	577	AC	ATTACCGCTGCGCACGCTACGGCCCTCCACACGGGGGCGCAATTGCGAGCCCTCCAG		636	
Db	214	Ala	PheProLeuArgAlaGlyGly	-----	AsnGlyGln	226
QY	637	TAT	TGGCCCTTTCTTCGCAATCTCTATAATTGGAAAACTAAACCATCCCCCTTTCTCG		696	
Db	227	Tyr	TrpProPheSerSerSerAspLeuTyrAsnTrpLysAsnAsnAsnProSerPheSer		246	
QY	697	GAG	ATCCCCAACCGCTCACGGGTGTGTGAGTCCCTTATGTCTCTCACAGCTACT		756	
Db	247	Glu	AspProGlyLysLeuThrAlaLeuIleGluSerValLeuIleThrHisGlnProThr		266	
QY	757	TGG	GATGATTCTCAACAGCTGCTCAGACACTCTTCACACCGAGGACGAGAGAGAATT		816	
Db	267	Trp	AspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgVal		286	
QY	817	CT	ATTAGAGGCTAGAAAAAATGTTCTCGGCGCGACCGGGCCACCCACCGCGGTGCAAAAT		876	
Db	287	Leu	LeuGluAlaArgLysAlaValArgLysAspAspGlyArgProThrGlnLeuProAsn		306	
QY	877	GAG	ATTGACATGGGATTTCCTTAATCTGCCCCCGGTGGGACTACAAACACCGCTGAAGGT		936	
Db	307	Glu	ValAspAlaAlaPheProLeuGluArgProAspTrpAspTyrThrThrGlnAlaGly		326	
QY	937	AGG	GAGACTTGAAATCTATCGCAGGCTCTGTGGCGGTCTCCGGGGCGCCTCAGA		996	
Db	327	Arg	AsnHisLeuValHisTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGlyArg		346	
QY	997	CGG	CCCACTAATTGGCTTAAGTAAGAGAAGTGCAGGAGCGACCGAATGAACCCCTCT		1056	
Db	347	Ser	ProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer		366	
QY	1057	GT	TTTTCTTGAGGCTCTTGGAGCCTTCAGGCGGTACACCCCTTTTGATCCCACTCA		1116	
Db	367	Ala	PheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp		386	
QY	1117	GAG	CCCCAAAAGCCTCAGTGGCTTTGGCCTTTATAGACAGTACAGCTTCGGATATTAGA		1176	
Db	387	Pro	GlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly		406	
QY	1177	AAG	AGCTTCAGAGACTTGAAGGCTTACAGAGGCTGAGTTACGTGACTAGTGAAGAG		1236	
Db	407	Arg	LysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu		426	
QY	1237	GC	AGAAAGTATATTACAAAGGGGACAGAGAAGAAAGGGACAAAGAAAGAGAGA		1296	
Db	427	Ala	GluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgIleArgArg		446	
QY	1297	GAA	GACAGGGAAGGGAGGAAGA	-----	CGT	1323
Db	447	Glu	ThrGluGluLysGluArgArgArgThrGluAspGluGlnLysGluLysGluArg		466	
QY	1324	AAT	AAACCGCAAGAGAGAAGATTTTGACATAAGATCTTGCTGCGAGTGGTTGAAGGGAAGC		1383	
Db	467	Asp	ArgArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys		486	
QY	1384	AAT	ACGGAAGAGAGAGAGAGATTTTATAGGAAATATAGGTACAGCCCTACAGACCTCAGGGAAC		1443	

Qy	1324	AATAAACGGCAAGAGAGAATTTCGCTAAGACTCTGGCTGCAGTGGTTGAAGCGAAAAGC	1383
Db		##### ::::##### ::::##### ::::##### ::::	
Db	467	AspArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys	486
Qy	1384	AATACGGAAAGAGAGAGAGATTTTAGGAAAAATTAGGTCCAGGCCCTAGCACAGTCAGGCAAC	1443
Db		##### ::::##### ::::##### ::::##### ::::	
Db	487	Gln-----AspArgGlnGlyGlyGlu	493
Qy	1444	CTGGGCAATAGGACCCCCTCGACAAGGACC AATGTGCATATTGTAAGAAAGAGGACAC	1503
Db		##### ::::##### ::::##### ::::##### ::::	
Db	494	--ArgArgArgSerGlnLeuAspArgAspGlnCysAlaTrpCyLysGluLysGlyHis	512
Qy	1504	TGGCGCAGGAAC TGCCCCCNAGAG-----GGAACAACAGGACCAAG-----	1545
Db		##### ::::##### ::::##### ::::##### ::::	
Db	513	TrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArgProGlnThrSer	532
Qy	1546	ATCCTAGCTCTAGAAGAA	1563
Db		##### ::::##### ::::##### ::::##### ::::	
Db	533	LeuLeuThrLeuAspAsp	538
 RESULT 8			
US-09-554-572-26			
; Sequence 26, Application US/09554572			
; Patent No. 6573091			
; GENERAL INFORMATION:			
; APPLICANT: NATURE TECHNOLOGY, INC.			
; TITLE OF INVENTION: CHIMERIC VIRAL PACKAGING SIGNAL WITHOUT GAG GENE			
; TITLE OF INVENTION: SEQUENCES			
; FILE REFERENCE: 228.00030201			
; CURRENT APPLICATION NUMBER: US/09/554,572			
; CURRENT FILING DATE: 2000-09-18			
; NUMBER OF SEQ ID NOS: 26			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 26			
; LENGTH: 1312			
; TYPE: PRT			
; ORGANISM: Murine leukemia virus			
US-09-554-572-26			

Alignment Scores:	
Pred. No.:	4.29e-119
Score:	1380.00
Length:	1312
Matches:	283
Percent Similarity:	68
Best Local Similarity:	Conservative: 68
Query Match:	Mismatches: 141
DB:	Indels: 74
	Gaps: 13
	2

US-10-723-552-3 COPY 585 2156 (1-1572) x US-09-554-572-26 (1-1312)

QY	1	ATGGGACAGACGGTCAGCACCCCTCTTAGTTTGACTCTCGACCAATTGGACTGAAGTTAA	60
Db	111	MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu	130
QY	61	TCGAGGGCTCATAATTGTCAGTTTTCAGTTAAGAAGGACCTTTGGCAGACTTTCTGTGC	120
Db	131	ArgIleAlaHisAsnGlnSerValAspValLysIysArgTrpValThrPheCysSer	150
QY	121	TCGTGAATGCCGACACTTCGATGTTTCGATGCGCATCAGGGGACCTTTAATCTCGAATT	180
Db	151	AlaGluIrrProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu	170
QY	181	ATCTCGGCTGTGTTAAAGCAGTTATTATTTTCAGACTCGACCCGGCTCTCATCCCGACTCAGGAG	240
Db	171	IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal	190
QY	241	CCCTATATCTTTCAGTCGACAGATTGTCAGAGACTCTCCGCCATGGGTAAACCATGG	300
Db	191	ProTyrlleValThrTrpGluAlaPheAspProProProTrpValLysProPhe	210
QY	301	CTGATATAGCCAGAAAGCCAGGTCCCGGAATTTCTGGCTCTTGGAGAGAAAAACAAC	360
Db	211	ValHis---ProLysProProProLeuProProSerAlaProSerLeuProLeuGlu	229

361 TCGGCTGAAAAAGTCAAGCCCTCTCTCCTCATATCTACTACCCGAGATTGAGAGGCCA---CCG 417

230 ProProArgSerThrProProArgSerLeuThrProAlaLeuThrProSerLeuGly 249

418 GCTTGGCCGGGAACCCAA-:::|||||:::-----TCT 438

250 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuLeuAspLeuLeuThr 269

439 GTTCCCCCACCCCTTATCTGGCACAGGTCGCGGAGGAGCCCTTTGCCCTCTCT--- 495

270 GluAspProProProTyr-----ArgAspProArgProProProSer 283

496 -----GGAGCTCGCGCGTGGAGGGACCT----- 519

284 AspArgAspGlyAsnGlyGlyAlaThrProAlaGlyGluAlaProAspProSerPro 303

520 ---GCTGAGGAGCTCGAGCGGAGGGCGGCCACCCCGGAGCGGACACAGACAGATCGCG 576

304 MetAlaSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrThrSerGln 323

577 ACATTACCGCTCGCACGTACGGCCCTCCACACCGGGGGCCAAATTGACGCCCTCCAG 636

324 AlaPheProLeuLeuAlaGlyGly-----AsnGlyGln-----LeuGln 336

637 TATTGGCCCTTTCTTCGACAGATCTTATAATTGGAATACTAACCATCCCTTTCTCG 696

337 TyrTrpProPheSerSerSerAspLeuTyrAsnTrpLysAsnAsnAsnProSerPheSer 356

697 GAGATCCCCAAACGCTCACGGGGTGTGTGAGTCCCTTATGTCTCTCACGAGCTACT 756

357 GluAspProGlyLysLeuThrAlaLeuLeuGluSerValLeuLeuThrHisGlnProThr 376

757 TGGGATGATTCTCAACAGCTCTCTGCAGACACTCTTTCACACCGAGGAGCGAGAGAATT 816

377 TrpAspAspCysGlnGlnLeuLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgVal 396

817 CTATTAGAGGCTAGAAAAAATGTTCTCTGGGCGGACGCGGACCCACGCGGTGCAAAAT 876

397 LeuLeuGluAlaAaGlyLysAlaValArgGlyAspAspGlyArgProThrGlnLeuProAsn 416

877 GAGATTGACATGGGATTTCCTTTAACTCGCCCGGTGGGAGCTACAAACGCGCTGAAGT 936

417 GluValAspAlaAlaPheProLeuGluArgProAspTrpAspTyrThrThrGlnAlaGly 436

937 AGGAGAGCTTGAAAATCTATCGCCAGCTCTGTGTGGCGGTCTCCGGGGCGCTCACA 996

437 ArgAsnHisLeuValHisTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 456

997 CGGCCCACTAATTTGGCTTAAGGTAAGAGAGTGTATGACGAGGACCGAATGAACCCCTCT 1055

457 SerProThrAsnLeuAlaLysValLysGlyLeuThrGlnGlyProAsnGlnSerProSer 476

1057 GTTTTTCTTGAGAGGCTTTGGAGACCTTCAGGCGGTACACCCCTTTTGATCCCCACCTCA 1111

477 AlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp 496

1117 GAGCCCCAAAAGCCTCAGTGGCTTTGGCCCTTTATAGGACAGTACAGCCTTGATATTAGA 1177

497 ProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly 516

1177 AAGAAGCTTCAGAGACTGGAAGGTTACAGGAGGCTGAGTTACGTGATCTAGTGAAGAG 1231

517 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 536

1237 GCAGAGAAAGTATATTACAAAGGGGACAGAGAAGAAAGGGGAACCAAGAAAAAGAGAGA 1299

537 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgIleArgArg 556

1297 GAAAGACAGAAAGGGGAAGA-----CGT 1321

557 GluThrGluGluLysGluGluArgArgArgThrGluAspGluGlnLysGluLysGluArg 576


```

Db 455 ArgArgThrGluAspGluGlnLysGluLysGluArgAspArgArgHisArgGluMet 474
QY 1348 ACTAAGATCTTGGCTGCAAGTGTTCGAAGGAAAGCAATACGAAAGAGAGAGAGATTTT 1407
Db 475 SerArgLeuLeuAlaThrValSerGlyGlnArgGlnAspArgGlnGlyGluArg 494
QY 1408 AGGAAATATTAGTCAAGCCCTAGACAGTCAGGGAACCTGGGCAATAGGACCCCACTCGAC 1467
Db 495 Arg-----ArgSerGlnLeuAsp 500
QY 1468 AAGGACCAATGTCATATTGTAAGAAAGAGACACTGGGCAAGAACTGCCCAAG--- 1524
Db 501 CysAspGlnCysThrTy-CysGluGlnGlyHisTrpAlaLysAspCysProArgArg 520
QY 1525 ---AAGGAAACAAAGGACCAAGG-----ATCCTAGCTCTAGAAGAA 1563
Db 521 ProArgGlyProArgGlyProArgProGlnThrSerLeuLeuThrLeuAspAsp 538

RESULT 10
US-08-850-961-12
; Sequence 12, Application US/08850961
; Patent No. 6013517
; GENERAL INFORMATION:
; APPLICANT: Respass, James G.
; APPLICANT: De Polo, Nicholas J.
; APPLICANT: Chada, Sunil
; APPLICANT: Sauter, Sybille
; APPLICANT: Bodner, Mordechai
; APPLICANT: Driver, David A.
; TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation, Intellectual Property - R440
; STREET: P.O. Box 8097
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,961
; FILING DATE: 05-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kruse, No. 6013517man J.
; REGISTRATION NUMBER: 35,235
; REFERENCE/DOCKET NUMBER: 930049.424C4 / 1147.005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-3520
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-850-961-12

Alignment Scores:
Pred. No.: 6.37e-29 Length: 140
Score: 403.50 Matches: 76
Percent Similarity: 61.1% Conservative: 15
Best Local Similarity: 51.0% Mismatches: 35
Query Match: 14.2% Indels: 23
DB: 2 Gaps: 2

US-10-723-552-3_COPY_585_2156 (1-1572) x US-08-850-961-12 (1-140)

```

```

QY 1 ATGGGACAGACGGTCAGACACCCCTCTTAGTTTGGACTCTCGACCATTTGACTGCAAGTTAAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
QY 61 TCCAGGGCTCATATTTGTCTAGTTTCAGGTTAAGAGAGGACCTTGGCAGACTTTCTGTGTC 120
Db 21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgArgTrpValThrPheCysSer 40
QY 121 TCTCAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTCAGATT 180
Db 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
QY 181 ATCTGCTGTTTAAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCGATCAGGAG 240
Db 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
QY 241 CCTATATCTCTAGTGGCAAGATTTCGACAGAGATCTCCGCCATGGTTAAACCATGG 300
Db 81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProTrpValLysProPhe 100
QY 301 CTGAATAAGCCAAAGAACCCAGGTCCTCCGAATTCCTGGCTCTTGGAGAGAAAAACAACAC 360
Db 101 ValHisPro----- 103
QY 361 TCGCTCAAAAAGTCAAGCCCTCTCTCATATCTACCCGAGATTGAGGAGCCACCGGCT 420
Db 104 -----LysProProProLeuProProSerAlaProSerLeuProLeu 118
QY 421 TGGCCGAACCCCAATCTGTTCCCCCA 447
Db 119 GluPro---ProArgSerThrProPro 126

RESULT 11
US-09-479-776-12
; Sequence 12, Application US/09479776
; Patent No. 633195
; GENERAL INFORMATION:
; APPLICANT: Respass, James G.
; APPLICANT: De Polo, Nicholas J.
; APPLICANT: Chada, Sunil
; APPLICANT: Sauter, Sybille
; APPLICANT: Bodner, Mordechai
; APPLICANT: Driver, David A.
; TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: INTELLECTUAL PROPERTY-R440
; P.O. BOX 8097
; CITY: EMERYVILLE
; STATE: CA
; COUNTRY: USA
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479,776
; FILING DATE: 07-Jan-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: KRUSE, NORMAN J.
; REGISTRATION NUMBER: 35,235
; REFERENCE/DOCKET NUMBER: 930049.424C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid

```

```
;
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-479-776-12

Alignment Scores:
Pred. No.: 6,37e-29 Length: 140
Score: 403.50 Matches: 76
Percent Similarity: 61.1% Conservative: 15
Best Local Similarity: 51.0% Mismatches: 35
Query Match: 14.2% Indels: 23
DB: 2 Gaps: 2

US-10-723-552-3_COPY_585_2156 (1-1572) x US-09-479-776-12 (1-140)
QY 1 ATGGGACAGAGCGGTGACGACCCCTTTAGTTGACTCTCGACCATTCGACTGGAAGTTAAA 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
QY 61 TCCAGGGCTCATATTTCTCAGTTTCAGGTTAAGAGGGACCTTGGCAGACTTCTGTGTC 120
DB 21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgA-gtrpValThrPheCysSer 40
QY 121 TCTGAATGCGCGACATTCGATGTTGATGATGATGATGATGATGATGATGATGATGAT 180
DB 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
QY 181 ATCTGGCTGTTAAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCCGATCAGAG 240
DB 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
QY 241 CCCTATATCCTTACGTGCAAGATTTGGCAGAGGATCTCCGCCATGGGTTAAACCATGG 300
DB 81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProProTrpValLysProPhe 100
QY 301 CTGAATAAGCCAAAGAACCCAGGTCCCGCAATTTCTGGCTCTTGGGAGAGAAAAACAAAC 360
DB 101 ValHisPro----- 103
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGAGGCCACCGCT 420
DB 104 -----LysProProProProProProProProProProProProProProPro 118
QY 421 TGGCGGAACCCCAATCTGTTCCCA 447
DB 119 GluPro----ProArgSerThrProPro 126
```

RESULT 12

```
; Sequence 10, Application US/08850961
; Patent No. 6013517
; GENERAL INFORMATION:
; APPLICANT: Respess, James G.
; APPLICANT: De Polo, Nicholas J.
; APPLICANT: Chada, Sunil
; APPLICANT: Sauter, Sybille
; APPLICANT: Bodner, Mordechai
; APPLICANT: Driver, David A.
; TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation, Intellectual Property - R440
; STREET: P.O. Box 8097
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
;
;
; APPLICATION NUMBER: US/08/850,961
; FILING DATE: 05-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kruse, No. 6013517man J.
; REGISTRATION NUMBER: 35,235
; REFERENCE/DOCKET NUMBER: 930049.424C4 / 1147.005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-3520
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-850-961-10

Alignment Scores:
Pred. No.: 6,39e-29 Length: 141
Score: 403.50 Matches: 76
Percent Similarity: 61.1% Conservative: 15
Best Local Similarity: 51.0% Mismatches: 35
Query Match: 14.2% Indels: 23
DB: 2 Gaps: 2

US-10-723-552-3_COPY_585_2156 (1-1572) x US-08-850-961-10 (1-141)
QY 1 ATGGGACAGAGCGGTGACGACCCCTTTAGTTGACTCTCGACCATTCGACTGGAAGTTAAA 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
QY 61 TCCAGGGCTCATATTTCTCAGTTTCAGGTTAAGAGGGACCTTGGCAGACTTCTGTGTC 120
DB 21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgA-gtrpValThrPheCysSer 40
QY 121 TCTGAATGCGCGACATTCGATGTTGATGATGATGATGATGATGATGATGATGAT 180
DB 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
QY 181 ATCTGGCTGTTAAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCCGATCAGAG 240
DB 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
QY 241 CCCTATATCCTTACGTGCAAGATTTGGCAGAGGATCTCCGCCATGGGTTAAACCATGG 300
DB 81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProProTrpValLysProPhe 100
QY 301 CTGAATAAGCCAAAGAACCCAGGTCCCGCAATTTCTGGCTCTTGGGAGAGAAAAACAAAC 360
DB 101 ValHisPro----- 103
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGAGGCCACCGCT 420
DB 104 -----LysProProProProProProProProProProProProProProPro 118
QY 421 TGGCGGAACCCCAATCTGTTCCCA 447
DB 119 GluPro----ProArgSerThrProPro 126
```

RESULT 13

```
; Sequence 10, Application US/09479776
; Patent No. 6333195
; GENERAL INFORMATION:
; APPLICANT: Respess, James G.
; APPLICANT: De Polo, Nicholas J.
; APPLICANT: Chada, Sunil
; APPLICANT: Sauter, Sybille
; APPLICANT: Bodner, Mordechai
; APPLICANT: Driver, David A.
; TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
; NUMBER OF SEQUENCES: 45
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: INTELLECTUAL PROPERTY-R440
P.O. BOX 8097
CITY: EMERYVILLE
STATE: CA
COUNTRY: USA
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, V
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/479,776
FILING DATE: 07-Jan-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: KRUSE, NORMAN J.
REGISTRATION NUMBER: 35,235
REFERENCE/DOCKET NUMBER: 930049.42
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)622-4900
TELEFAX: (206)682-6031
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-479-776-10

```

```

Db      119 Glupro---ProArgSerThrProPro 126
RESULT 14
US-09-719-554-58
; Sequence 58, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: RIEGER, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE 1
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 58
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (86)..(86)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (166)..(166)

```

Alignment Scores:	6.39e-29	Length:	141
Pred. No.:	403.50	Matches:	76
Score:	403.50	Conservative:	15
Percent Similarity:	61.1%	Mismatches:	35
Best Local Similarity:	51.0%	Indels:	23
Query Match:	14.2%	Gaps:	2
DB:	2		

US-10-723-552-3 COPY 585 2156 (1-1572) x US-09-479-776-10 (1-141)

QY	1	ATGGCAGCAGCGGTGACGACCCCTCTTAGTTTGACTCTCGACCAATTGCACTGAAGTTAA	60
Db	1	MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpValThrPheCysValGlu	20
QY	61	TCGAGGGCTCATAAATTTGCTAGTTTACAGTTTAAGAAGGGACCTTGGCAGACTTCTGTGTC	120
Db	21	ArgIleAlaHisAsnGlnSerValAspValIleValysArgTrpValThrPheCysSer	40
QY	121	TCTGAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAATCTCGAGATT	180
Db	41	AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu	60
QY	181	ATCTCGGCTGTTAAACAGATTATTTTTCAGACTGGACCCGGCTCTCATCCCGACTCAGGAG	240
Db	61	IleThrGlnValIleValIleValPheSerProGlyProHisGlyHisProAspGlnVal	80
QY	241	CCCTATATCTTACCTGGCGAAGATTGGCAGAGATCTCCGCCATGGGTTAAACCATGG	300
Db	81	ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProTrpValIleValysProPhe	100
QY	301	CTGAATAAGCCAGAAACCCAGTCCCCGAATTTCTGGCTCTTGGAGAGAAAACCAACAC	360
Db	101	ValHisPro-----	103
QY	361	TGGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCACCGGCT	420
Db	104	-----LysProProProLeuProSerAlaProSerLeuProLeu	118
QY	421	TGGCCGGAACCCCAATCTGTTCCCCCA	447

Alignment Scores:		
Pred. No.:	5.6e-17	Length:
Score:	277.00	Matches:
Percent Similarity:	39.6%	Conservative:
Best Local Similarity:	26.9%	Mismatches:
Query Match:	9.7%	Indels:
DB:	2	Gaps:
		420

US-10-723-552-3 COPY 585 2156 (1-1572) x US-09-719-554-58 (1-420)

QY 517 CCTGCTCAGGGACTCGGAGCCGGAGGGCGCCACCCCGAGCGGACAGACGAGATC-- 573


```

Db      27  GlnThr-----LeuArgLysLysArgPhePheCysSerThr 40
Qy      124  GAATGCCGACATTCGATGTT-----GGATGGCCATCAGAGGGGACCTTTAAT 171
Db      41  AlaTyrProGlnTyrProLeuGlnGlyArgGluThrTrpLeuProGluGlySerIleAsn 60
Qy      172  TCTGAGATTATCTGGCTGTAAACAGATTATTTTCAGACTGGACCGGCTCTCATCCC 231
Db      61  TyrAsnIleLeuGlnLeuAspLeuPheCysArgLysGluGlyLysTrpSer----- 78
Qy      232  GATCAGGAGCCCTATATCTTACGTGGCAAGATTTCGGCAGAGGATCCTCCGCCATGGGTT 291
Db      79  ---GluValProTyrValGlnThrPheSerLeuArgAsp----- 91
Qy      292  AAACCATGGCTGAATAAGCAAGAAAGCCAGGTCCCGAATTCGTGGCTCTTGGAGAGAA 351
Db      91  ----- 91
Qy      352  AACAAACACTCGGCTGAAGAAAGTCAAGCCCTCTCTCATATCTACCCGAGATTGAGGAG 411
Db      92  AsnSerGlnLeuCysLysLysCysGlyLeuCysProThrGlySerPro-----GlnSer 109
Qy      412  CCACCGGCTTGGCGGACCCCAATCTGTCTCCCCACCCCTTATCTGCGCACGGGTGCC 471
Db      110  ProProTyrPro-----SerValProSerPro----- 119
Qy      472  GCGAGGGGACCTTTGCCCCCTCTGGAGCTCCGGCGGTGGAGGACCTGCTGCGGGACT 531
Db      120  -----ThrProSerSerThrAsnLysAspProProLeuThrGlnThr 133
Qy      532  CGGAGCCGG-----AGGGGGCCACCCCGGAGCGGACAGAGATCGCACATTA 582
Db      134  ValGlnLysGluIleAspLysGlyValAsnAsnGlu-----ProLysSerAlaAsnIle 151
Qy      583  CCGCTGGCGACGTACGGCCCTCCACACCGGGGGGCCAATTGCGAGCCCTCCAGTAT--- 639
Db      152  ProArgLeuCysProLeuGlnAlaValArgGlyGlyGluPheGlyProAlaValPro 171
Qy      640  TGGCCCTTTCTCTTCAGATCTCTATAATTGGAATACTAACCATCCCTTTCTCGGAG 699
Db      172  ValProPheSerLeuSerAspLeuLysGlnIleLysIleAspLeuGlyLysPheSerAsp 191
Qy      700  GATCCCCACCCCTCACGGGCTTGGTGAGTCCCTTATGTTCTCTACCCAGCTACTTGG 759
Db      192  AsnProAspGlyTyrIleAspValLeuGlnGlyLeuGlyGlnSerPheAspLeuThrTrp 211
Qy      760  GATGATTGTCAACAGCTGCTGCAGACACTCTTCACACCGGAGGAGGAGAGAGAAATCTA 819
Db      212  ArgAspIleMetLeuLeuLeuAsnGlnThrLeuThrProAsnGluArgSerAlaVal 231
Qy      820  TTAGAGGCTAGAAA-----AATGTTCCTGGGGCGGACGGCGGCCACCG 864
Db      232  ThrAlaAlaArgGluPheGlyAspLeuTrpTyrLeuSerGlnAlaAsnAsnArgMetThr 251
Qy      865  CGGTTGCAAAATGATGATGATCGGA-----TTTCCCTTAATCGCCCGGTTGG 915
Db      252  ThrGluGluArgThrThrProThrGlyGlnGlnAlaValProSerValAspProHisTrp 271
Qy      916  GACTACACACCGCTGAGGTAGGAGAGCTTGAATAATCTATCCCGAGGCTCTGCTGGCG 975
Db      272  AspThrGluSerGluHisGlyAspTrpCysHisLysHisLeuLeuThrCysValLeuGlu 291
Qy      976  GGTCTCCGGGCGCTCAAGACGGCCCACTAATTTGGCTAAGGTAAGGAAGTATGTCAG 1035
Db      292  GlyLeuArgLysThrArgLysLysProMetAsnTyrSerMetSerThrIleThrGln 311
Qy      1036  GGACCGAATGAACCCCTCTGTTTCTTGAGAGGCTCTTGGAAAGCTTCAGCGGTAC 1095
Db      312  GlyLysGluGluAsnLeuThrAlaPheLeuAspArgLeuArgGluAlaLeuArgLysHis 331
Qy      1096  ACCCTTTTGTATCCACTCGAGGGCCCAAAAGCCCTCAGTGGCTTTGGCTTTATAGGA 1155
Db      332  ThrSerLeuSerProAspSerIleGluGlyGlnIleLeuLysAspLysPheIleThr 351

```

Qy 1156 CAGTCAGCCTTGGATATTAGAAAGAGCTTCAGAGACTTGAAGGGTTACAGGAGGCT 1212
 Db 352 GlnSerAlaAlaAspIleArgLysAsnPheLysSerLeuProLysLeuAlaAla 370

Search completed: February 14, 2006, 16:15:34
 Job time : 68.6237 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - protein search, using frame_plus_n2p model
Run on: February 14, 2006, 14:51:21 ; Search time 47.4857 Seconds
(without alignments)
3295.451 Million cell updates/sec
Title: US-10-723-552-3
Perfect score: 14636
Sequence: 1 GCGTGGTGACGACTGTGGG.....CTGTTTGCATCAAAAAAAA 8132
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 566832
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10723552/runat.14022006.125143.12937/app.query.fasta_1
-DB=PIR -QPMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN 1.1.107 @runat.14022006.125143.12937 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5237	35.8	1784	2 T10532	gag-pol polyprotei
2	4316.5	29.5	1165	1 GNLLJGL	HIV-1 retropepsin
3	3999	27.3	1196	1 GNMVGV	HIV-1 retropepsin
4	3922.5	27.3	1204	2 S35475	pol polyprotei
5	3980.5	27.2	1204	2 S70393	pol polyprotei
6	3973.5	27.1	1199	1 GNMVIM	HIV-1 retropepsin
7	3947	27.0	1196	1 GNMVRV	HIV-1 retropepsin
8	3940.5	26.9	1189	1 GNMVM7	HIV-1 retropepsin
9	3536.5	24.2	1046	1 GNMVCE	pol polyprotei
10	2639.5	18.0	843	1 GNMVK	pol polyprotei
11	1862.5	12.7	559	2 A46311	pol polyprotei
12	1725.5	11.8	476	2 S04842	pol polyprotei
13	1681	11.5	581	2 A42743	pol polyprotei
14	1677.5	11.5	520	1 FOLJGL	gag polyprotei

15	1643	11.2	1751	2 T09394	gag-pro-pol polypr
16	1637.5	11.2	512	1 FOMVGS	gag polyprotei
17	1505.5	10.3	667	1 VCLJGL	env polyprotei pr
18	1477.5	10.1	537	1 FOMVM7	gag polyprotei
19	1383.5	9.5	538	2 S35474	gag polyprotei
20	1381.5	9.4	538	2 S70394	gag polyprotei
21	1379.5	9.4	538	1 FOMVM	gag polyprotei
22	1376.5	9.4	537	1 FOMVIM	gag polyprotei
23	1371	9.4	537	1 FOMVGV	gag polyprotei
24	1364	9.3	537	1 FOMVRV	gag polyprotei
25	1359.5	9.3	540	1 FOMVHL	gag polyprotei
26	1357	9.3	537	1 FOMVMB	gag polyprotei
27	1351	9.2	536	2 S70395	env polyprotei
28	1347.5	9.2	536	1 FOMVME	gag polyprotei
29	1347	9.2	676	2 T01381	env protein - muri
30	1345	9.2	676	1 VCMVPV	env polyprotei pr
31	1344	9.2	665	1 VCMVKA	env polyprotei pr
32	1339	9.1	665	1 VCMVVR	env polyprotei pr
33	1339	9.1	669	1 VCMVEK	env polyprotei
34	1335	9.1	669	2 A46511	envelope protein -
35	1335	9.1	689	2 B43491	env polyprotei
36	1332.5	9.1	688	2 A43491	env polyprotei
37	1326	9.1	529	1 FOMVDD	gag polyprotei
38	1320	9.0	665	1 VCMVEM	env polyprotei
39	1313	9.0	666	1 VCMVHL	env polyprotei pr
40	1312	9.0	532	1 FOMVVB	gag polyprotei
41	1290	8.8	640	1 VCMVRV	env polyprotei pr
42	1288	8.8	661	1 VCMVCB	env polyprotei
43	1276	8.7	642	1 VCMVFG	env polyprotei
44	1273	8.7	642	2 T10533	env polyprotei pr
45	1268.5	8.7	536	1 FOMVMD	gag polyprotei

ALIGNMENTS

RESULT 1

T10532

gag-pol polyprotei precursor - feline leukemia virus (strain FeLV-FAIDS)
N;Contains: core protein p10; core protein p12; core protein p15; core protein p27; end
C;Species: feline leukemia virus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10532

J;Donahue, P.R.; Hoover, E.A.; Beltz, G.A.; Riedel, N.; Hirsch, V.M.; Overbaugh, J.; Mu
R. Virol. 62, 722-731, 1988
A;Title: Strong sequence conservation among horizontally transmissible, minimally patho
A;Reference number: Z17078; MUID:88119207; PMID:2828667
A;Accession: T10532
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-1784 <DON>
A;Cross-references: UNIPROT:Q85521; UNIPARC:UPI000010C1P5; EMBL:M18247; NID:G3233904; PI
C;Superfamily: pol polyprotei
C;Keywords: hydrolase; nucleotidyltransferase; polyprotei; reverse transcriptase
F;75-576/Product: gag polyprotei #status predicted <GAG>
F;702-271/Product: core protein p15 #status predicted <P15>
F;202-271/Product: core protein p12 #status predicted <P12>
F;272-519/Product: core protein p27 #status predicted <P27>
F;520-576/Product: core protein p10 #status predicted <P10>
F;577-1784/Product: pol polyprotei #status predicted <POL>
F;577-701/Product: proteinase #status predicted <PTN>
F;702-1368/Product: RNA-directed DNA polymerase (EC 2.7.7.49) #status predicted <REV>
F;1369-1784/Product: endonuclease (EC 3.1.1.-) #status predicted <EDC>

Alignment Scores:
Pred. No.: 2,74e-314 Length: 1784
Score: 5237.00 Matches: 1028
Percent Similarity: 70.3% Conservative: 232
Best Local Similarity: 57.3% Mismatches: 416
Query Match: 35.8% Indels: 117
DB: 2 Gaps: 27

US-10-723-552-3 (1-8132) x T10532 (1-1784)

```
Qy 512 TCAGTGTGCTGGATCTGTGGTTCCTGTTTCTGTGTGCTTCTGCTGTGTGCTTGT 571
Db ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
47 SerArgAlaAlaArgLeuValIlePheCysLeuValAlaSerPheLeuVal---ProCys 65
Qy 572 CTACAGTTTTT-----AATATGGGACAGACGGTGCAGCCCTCTTAGTTT 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 Leu-ThrPheLeuIleAlaGluThrValMetGlyGlnThrIleThrThrProLeuSerLe 85
Qy 617 GACTCTCGACATTGGACTGAAGTTAAATCAGGGCTCATATTTGTTCAGTTTCAGGTTAA 676
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
85 uThrLeuAspHisTrpSerGluValArgAlaArgAlaHisAsnGlnGlyValGluValArg 105
Qy 677 GAAGGACCTTGGCAGACTTCTGTCTCTCGATGCGCCGACATTCGATGTTGATGCGC 736
Db :||| ||| ||| ||| :||| ||| ||| ||| :||| ||| ||| ||| :||| ||| ||| |||
105 GlyLeuLysTrpIleThrLeuCysGluAlaGluTrpValMetMetAsnValGlyTrpPr 125
Qy 737 ATCAGAGGGGACCTTAAATCTCGAGATTATCTCGCTGTTAAACAGTATTATTTTCAGAC 796
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
125 oArgGluGlyThrPheSerLeuAspAsnIleSerGlnValGluLysIlePheAlaPr 145
Qy 797 TGGACCCGGCTCTCATCCGATCAGGAGCCCTATATCTTACGTGGCAAGATTGGCAGA 856
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
145 oGlyProTyrglyHisProAspGlnValProTyrglyIleThrThrTrpArgSerLeuAlaTh 165
Qy 857 GGATCTCCGCCATGGGTAAACCATCGCTGAATAAGCCAGAAAGCCAGTCCCGCGAAT 916
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 rAspProSerTrpValArgProPheLeuProProProlysThrProThrPro----- 183
Qy 917 TCTGGCTCTTGGAGAGAAAAACAAACATCTCGCTGAAAAAGTCAAGCCCTCTCCT--- 971
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
184 -----LeuProGlnProLeuSerProGlnProSerAl 194
Qy 972 -----CATATCTACCCCGAGATT-----GAGGAGCACCCGCG 1003
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
194 aProLeuThrSerSerLeuTyProValLeuProLysSerAspProProlysProProVa 214
Qy 1004 TTGGCCGGAACCCCAATCTGTTC-----CCACCCCTTA 1039
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 lLeuProProAspProSerSerProLeuIleAspLeuLeuThrGluGluProProProTy 234
Qy 1040 TCTGGCAGAGGTCCCGAGGAGCCCTTGGCCCTCTCTGGAGCTCCG-----GCGGT 1093
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 r-----ProGlyHisGlyPro-----ProProSerGlyProArgThrProTh 249
Qy 1094 GGAGGACCTCTCGAGGACTCGAGCCGAGGGCGGCCACCCCGAGCGGACAGACGA 1153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 rAlaSerProIleAlaSerArgLeuArgGluArgGluAsnPro-----AlaGluGl 267
Qy 1154 GATCGGACATTACCGTGGCAGCTACGGCCCTCCACACCGGGGGGCCAATTGCAGCC 1213
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
267 uSerGlnAlaLeuProLeuArg-----GluGlyProAsnAsnArgPr 281
Qy 1214 CCTCAGTATTGGCCCTTTCTCTCGAGATCTCTATAATTGGAATAACCAATCCCCC 1273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
281 o---GlnTyTrpProPheSerAlaSerAspLeuTyAsnTrpLysSerHisAsnProPr 300
Qy 1274 TTTCTCGAGGATCCCAACGCCCTCAGCGGTTGGTGGAGTCCCTTATGTTCTCTACCA 1333
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 oPheSerGlnAspProValAlaLeuThrAsnLeuIleGluSerIleLeuValThrHisGl 320
Qy 1334 GCCTACTGGGATGTTGCAACAGCTGTCGACACACTCTTCACACCGGAGGACGAGA 1393
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 nProThrTrpAspAspCysGlnGlnLeuLeuGlnAlaLeuLeuThrGlyGluGluArgGl 340
Qy 1394 GAGATTCTATTAGAGGCTAGAAAAATGTTCTTGGGGCCGACGGGGACCCACCGCGTT 1453
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
340 nArgValLeuLeuGluAlaArgLysGlnValProGlyGluAspGlyArgProThrGlnLe 360
Qy 1454 GCAAAATGAGATTGACATGGGATTTCCCTTAATCGCCCGCGGTGGGACTACACAGCG 1513
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 uProAsnValIleAspGluAlaPheProLeuLeuThrArgProAsnTrpAspPheArgTrPr 380
Qy 1514 TGAAGGTAGGAGAGCTTGAAATCTATCGCAGGCTCTGTGTGGGTCTCTCCGGGCGC 1573
```

```
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
380 oAlaGlyArgGluHisLeuArgLeuTyArgGlnLeuLeuAlaGlyLeuArgGlyAl 400
Qy 1574 CTCAGAAGCCGACCACTAATTTGGCTAAGTAAGAGAAGTATGACGGGACCGAATGAACC 1633
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 aAlaArgArgProThrAsnLeuAlaGlnValLysGlnValValGlnGlyLysGluGluTh 420
Qy 1634 CCCTCTGTTTTTCTGAGAGGCTCTTGGAGGCTTCAGGGCTTATAGGACAGTCAGCTTGGATCC 1693
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 rProAlaSerPheLeuGluArgLeuLysGluAlaTyArgMetTyThrProTyAspPr 440
Qy 1694 CACCTCAGAGGCCCAAAAGCCCTCAGTGGCTTGGCTTTATAGGACAGTCAGCTTGGGA 1753
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
440 oGluAspProGlyGlnAlaSerValIleLeuSerPheIleTyrglnSerSerProAs 460
Qy 1754 TATTAGAAAGAAGCTTCAGAGACTGGAAGGTTTACAGAGGCTGAGTTAGTATCTAGT 1813
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
460 pIleArgAsnLysLeuGlnArgLeuGluGlyLeuGlnGlyPheThrLeuSerAspLeuLe 480
Qy 1814 GAAGGAGCAGAGAAAGTATATTACAAAGGGGACAGAGAAGAAAGGGAACAAAGAA 1873
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
480 uLysGluAlaGluLysIleTyrglnLysArgGluThrProGluGluArgGluGluArg-- 499
Qy 1874 AGAGAGAGAAAGAGAGAGGAGGAGAAAGACGCTAATAAAGCGCAAGAGAATTTGAC 1933
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 -----LeuTrpGlnArgGlnGluGluArgAspLysArgHisLysGluMetTh 516
Qy 1934 TAAGATCTTGGCTCGAGTGTGTAAGGAAAGCAATACGGAAGAGAGAGA----- 1985
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
516 rLysValLeuAlaThrValValAlaGlnAsnArgAspLysAspArgGluGluSerLysLe 536
Qy 1986 ----GATTTTAGGAAATTAGTTCAGGCCCTTAGACAGTCAGGGAACCTGGGCAATAGGAC 2041
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
536 uGlyAspGlnArgLysIle----- 542
Qy 2042 CCACCTCGACAGACCAATGTGCATATTGTAAGAAAGAGGACACTTGAAGTGGAGGCGCAACC 2101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
543 -ProLeuGlyLysAspGlnCysAlaTyrglnLysGluLysGlyHisTrpValArgAspCy 562
Qy 2102 CCCCAAGAGGGAAACAAAGGACCAAGATCTTAGTCTTAGAAGAGATAAAGATTAGGG 2161
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
562 sProLysArgProArgLysLysProAlaAsnSerThrLeuLeuAsnLeuGluAsp***Gl 582
Qy 2162 GAGACGGGTTTCGAGCCCTCCGAGCCAGGTAACCTTGAAGTGGAGGCGCAACC 2221
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
582 uSerGlnGlyGlnAspProProGluProArgGlyThrLeuLysIleGlyGlnPr 602
Qy 2222 AGTTGAGTTCCTGTTGATACCGAGCGAAACATTCAAGTCTACTACAGCCATTAGGAAA 2281
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
602 ovalThrPheLeuValAspThrGlyAlaGlnHisSerValLeuThrArgProAspGlyPr 622
Qy 2282 ACTAAAGATAAAATCTCTGGGTGATGGTGCACAGGCGCAACACAGTATCATGAGAC 2341
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
622 oLeuSerAspArgThrAlaLeuValGlnGlyAlaThrGlySerLysAsnTyArgTrpTh 642
Qy 2342 TACCCGAGACAGTTCAGTCTGGAGTGGGAGTGAACCCACTCTCTTCTGGTCATACC 2401
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
642 rThrAspArgArgValGlnLeuAlaThrGlyLysValThrHisSerPheLeuTyValPr 662
Qy 2402 TGAGTGCACAGACCCCTCTTAGTAGAGACTTATTGACCAAGATGGGACACAAATTTTC 2461
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
662 oGluCysProTyrglnLeuLeuGlyArgAspLeuLeuThrLysLeuLysAlaGlnIleHi 682
Qy 2462 TTTTGAA---CAAGGGAAACAGAGTGTCTGCAATATAACAAACCTATCATCTGTGTGAC 2518
Db sPheThrGlyGluGlyAlaAsnValValGlyProArgGlyLeuProLeuGlnValLeuTh 702
Qy 2519 CCTCCATATGATGACCAATATCGACTATCTCTCCCTAGTAAAGCTGATCAAAATAT 2578
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
702 rLeuGlnLeuGluGluTyrglnLeuPheGluProGluSerThrGlnLysGlnGluMe 722
Qy 2579 ACAATCTCGTTGGAAACAGTTTCCCAAGGCTCGGCGAGAAACCGCAGGATGGGTTGGC 2638
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 722 tAspIleTrpLeuLysAsnPheProGlnAlaTrpAlaGluThrGlyGlyMetGlyMetal 742
Qy 2639 AAAGCAAGTTCCCCCAAGAGTTATTCAACTGAAGCCAGTGCACACAGCAGTGTCAAGTCAAG 2698
Db 742 aHisCysGlnAlaProValLeuIleGlnLeuLysAlaThrAlaThrProIleSerIleAr 762
Qy 2699 ACAGTACCCCTTGAGTAAGAAGCTCAAGAGGAATTCGGCCGCGATGTCGAAGATTAAAT 2758
Db 762 gGlnTyrProMetProHisGluAlaTyrGlnGlyIleLysProHisIleArgArgMetLe 782
Qy 2759 CCAACAGCGGATCTAGTCTCTGCTCCCAATCTCCCTGGAATACTCCCTGCTACCGGTTAG 2818
Db 782 uAspGlnGlyIleLeuLysProCysGlnSerProTrpAsnThrProLeuLeuProValIy 802
Qy 2819 AAAGCTGGAGCTAATGACTATCGACAGTACAGGACTTGAGAGAGGTCAATAAACGGGT 2878
Db 802 sLysProGlyThrLysAspTyrArgProValGlnAspLeuArgGluValAsnLysArgVa 822
Qy 2879 GCAGGATATACACCAACAGTCCGAACCTTATAACTCTTGCTGCTCTCCACCCCA 2938
Db 822 lGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerThrLeuProProSe 842
Qy 2939 ACGGAGCTGTATACAGTATTGGACTTAAAGGATGCTCTCTGCTGAGATTACACCC 2998
Db 842 rHisProTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisSe 862
Qy 2999 CACTAGCCAAACCTTTTCCCTTCGAATGGAGATCCAGAGTACGGGAAGAACCAGGCA 3058
Db 862 rGluSerGlnLeuLeuPheAlaPheGluTrpArgAspProGluIleGlyLeuSerGlyGI 882
Qy 3059 GCTCAGCTGACCGGACTGCCAGGGTTCAGAACTCCCGGACCACTTTTGACGAAGC 3118
Db 882 nLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAspGluAl 902
Qy 3119 CCTACACAGAGATGSCCAACTTCAGGATCCAAACCTCCAGTGACCTCCCTCCAGTA 3178
Db 902 aLeuHisSerAspLeuAlaAspPheArgValArgTyrProAlaLeuValLeuLeuGlnTy 922
Qy 3179 CGTGGATGACCTGCTTCGCGGGAGCCACAAACAGGACTGTAGAGGCGCAGGAGGC 3238
Db 922 rValAspAspLeuLeuLeuAlaAlaAlaThrArgThrGluCysLeuGluGlyThrLysAl 942
Qy 3239 ACTACTGCTGGAAATTTGCTGACCTAGCTACAGAGCTCTGCTAAGAGCCCGAGATTG 3298
Db 942 aLeuLeuGluThrLeuGlyAsnLysGlyTyrArgAlaSerAlaLysLysAlaGlnIleCy 962
Qy 3299 CAGGAGAGGTAACTACTTGGGTACACTTTGCGGAGCGGACGGATGGCTGACGGA 3358
Db 962 sLeuGlnGluValThrTyrLeuGlyTyrSerLeuLysAspGlyGlnArgTrpLeuThrLy 982
Qy 3359 GGCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACACAGCCAAACAAATGAGAGA 3418
Db 982 sAlaArgLysGluAlaIleLeuSerIleProValProLysAsnProArgGlnValArgGI 1002
Qy 3419 GTTTTGGGACAGCTGGATTGTCAGCTAGCTACAGAGCTCTGCTAAGAGCCCGAGATTG 3478
Db 1002 uPheLeuGlyThrAlaGlyTyrCysArgLeuTrpIleProGlyPheAlaGluLeuAlaI 1022
Qy 3479 CCCACTACCCGCTAACCAAGAAAGGGGAATTCCTCGGCTCCTGAGCACCAGAA 3538
Db 1022 aProLeuTyrProLeuThrArgProGlyThrLeuPheGlnTrpGlyThrGlnGlnLe 1042
Qy 3539 GGCATTGTGATGCTATCAAAAGCCCTGCTGAGCGACCTGCTCTCGCCCTCCCTGACGT 3598
Db 1042 uAlaPheGluAsnIleArgLysAlaLeuLeuSerSerProAlaLeuGlyLeuProAspIl 1062
Qy 3599 AACTAAACCTTTACCTTTATGTGATGAGCTTAAGGAGTAGCCCGGGAGATTTAAC 3658
Db 1062 eThrLysProPheGluLeuPheIleAspGluAsnSerGlyPheAlaLysGlyValLeuVa 1082
Qy 3659 CCNAACCTTAGGACCAAGAGACCTGCTGCTACCTGCTCAAGAACCTGATCTGCT 3718
Db 1082 lGlnLysLeuGlyProTrpLysArgProValAlaTyrLeuSerLysLysLeuAspThrVa 1102

Qy 3719 AGCAGTGGTGGCCCATATGCTGAAGGCTATCGAGCTGTGCGCATACTGCTCAAGGA 3778
Db 1102 lAlaSerGlyTrpProCysLeuArgMetValAlaAlaIleAlaIleLeuValLysAs 1122
Qy 3779 CGCTGACAAATTGACTTTGGGACAGAAATATACTGTAATAGCCCCCATGATTTGGAGAA 3838
Db 1122 pAlaGlyLysLeuThrLeuGlyGlnProLeuThrIleLeuThrSerHisProValGluAl 1142
Qy 3839 CATCGTTCCGCGCCCGCCAGACCGATGATGACCAACGCCCGCATGACCCACTATCAAG 3898
Db 1142 aLeuValArgGlnProProAsnLysTrpLeuSerAsnAlaArgMetThrHisThrGlnAl 1162
Qy 3899 CTGCTTCTC- --ACAGAGAGGGTCACTTTCGCTCCACAGCGCGCTCTCAACCTCTGCCAC 3955
Db 1162 aMetLeuLeuAspAlaGluArgValHisPheGlyProThrValIserLeuAsnProAlaTh 1182
Qy 3956 TCTTCTCGCTGAAGACTGATGAACCA- ----GTGACTCATGATTGCCACTCACTATT 4009
Db 1182 rLeuLeuProLeuProSerGlyLysProProArgLeuSerProAsp- ---- 1197
Qy 4010 GATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTCGGAGAGTCT 4069
Db 1198 -LeuAlaGluThrMetAlaGlnThrAspLeuThrAspGlnProLeuProAspAlaAspLe 1217
Qy 4070 AACCTGGTTCACTGACGGAAGCAGCTATGTGTGGGAAGGTAAGAGGATGCTCGGGCGGC 4129
Db 1217 uThrTrpTyrThrAspGlySerSerPheIleArgAsnGlyGluArgLysAlaGlyAlaAl 1237
Qy 4130 GGTGTGGAGCGGACCCGACAGATCTGGGCCAGCAGCTGCGGAGAGAACTTTCAGACA 4189
Db 1237 aValThrThrGluSerGluValIleTrpAlaAlaSerLeuProProGlyThrSerAlaGI 1257
Qy 4190 AAAGGTGAGCTCATGCGCTCAGCAAGCTTTGCGCTGCGCGGAGGGAATCCATAA 4249
Db 1257 nArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaLysGlyLysLeuTh 1277
Qy 4250 CATTTATACGACGACGAGGTATGCTTTCGACTGCACACGCTACATGCGGCGCATATAA 4309
Db 1277 rValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyGluIleTyrAr 1297
Qy 4310 ACNAAGGGGTGCTTACTCTACGAGGAGGAGGAATAAAGACAAAGAGAAATTTCTAAG 4369
Db 1297 gArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAsnGluIleLeuAl 1317
Qy 4370 CTTATTAGAGCGGTACATTTTACCAAAAGGCTAGCTATTATACATGCTCCGACATCA 4429
Db 1317 aLeuLeuGluAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHisGI 1337
Qy 4430 GAAAGCTAAAGATCTCATATCCAGAGGAAACAGATGCTGACCGGTTGCCAAGCAGGC 4489
Db 1337 nLysGlyAspSerProGlnAlaLysGlyAsnArgLeuAlaAspAspThrAlaLysLysAl 1357
Qy 4490 AGCC- ----CAGGGTGTAACTTCTGCT- ----ATAATAGAATGCCAA 4531
Db 1357 aAlaThrGluThrGlnSerSerLeuThrIleLeuProThrGluLeuIleGluGlyProLy 1377
Qy 4532 AGCCCCAGAAACCCAGACGACGATACACCTAGAGACTGCGCAGAGATATAAAGATAGA 4591
Db 1377 s- ----ArgProProTrpGluTyrAspAspSerAspLeuAspLeuValGlnLysLeuGI 1395
Qy 4592 CAGTTCTCTGAGACTCCGGAAGGACCTGTATACCTCAGATGGGAAGAAATCTCGCC 4651
Db 1395 uAlaHisTyrGluProLysArgGlyThr- --TrpGluTyrArgGlyLysThrIleMetPr 1414
Qy 4652 CCACAAAGAGGGTTAGAAATATGTCCACAGATACATGCTTAACCCACCTAGGAATAA 4711
Db 1414 oGluLysTyrAlaLysGluLeuIleSerHisLeuHisLysLeuThrHisLeuSerAlaAr 1434
Qy 4712 ACACCTGACGAGTTGGTCAGAACATCC- ----CCTTATCATGTTCT 4753
Db 1434 gLysMetLysThrLeuLeuGluArgGluThrGlyPheTyrLeuProAsnArgAspLe 1454

Db 220 LysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluIleAsnLysArgVal 239
QY 2880 CAGGATATACACCAACAGTCCCGAAACCTTATTAACCTCTGTGTGCTCTCCACCCCAA 2939
Db 240 GlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSerSerLeuProProSer 259
QY 2940 CGGAGCTGGTATACAGTATTGGACTTAAAGATGCCCTTCTCTGCTGAGATACACCCC 2999
Db 260 TyrThrTrpTyrSerValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisPro 279
QY 3000 ACTAGCCAAACACTTTTGGCTTCGAATGAGAGATCCAGTACCGGAAAGAACCGGGCAG 3059
Db 280 AsnSerGlnProLeuPheAlaPheGluTrpLysAspProGluLysGlyAsnThrGlyGln 299
QY 3060 CTACCTGGACCCGACTCCCGAAGGTTCAAGAACTCCCGACCATCTTTGACGAAGCC 3119
Db 300 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAspGluAla 319
QY 3120 CTACACAGACCTGGCCAACTTCAGGATCCAAACACCTCAGGTGACCTCTCCAGTAC 3179
Db 320 LeuHisArgAspLeuAlaProPheArgAlaLeuAsnProGlnValValLeuLeuGlnTyr 339
QY 3180 GTGGATGACCTGCTTCTGGCGGAGCCACCAACAGGACTGCTTAGAGCAGCAGAGCCA 3239
Db 340 ValAspAspLeuLeuValAlaAlaProThrTyrGluAspCysLysLysGlyThrGlnLys 359
QY 3240 CTACTGCTGCAATTGCTGACTAGGCTACAGAGCTCTGCTAAGAGGCCAGATTGTC 3299
Db 360 LeuLeuGlnLeuSerLysLeuGlyTyrArgValSerAlaLysLysAlaGlnLeuCys 379
QY 3300 AGGAGAGAGTAAACATCTTTGGGGTACAGTTTGGGGACCGGACGAGTGGCTGACGGAG 3359
Db 380 GlnArgGluValThrTyrLeuGlyTyrLeuLysGluGlyLysArgTrpLeuThrPro 399
QY 3360 GCACGGAGAACTGTAGTCCAGATACCGGCCCAACACACGACCAAAACAAATGAGAGAG 3419
Db 400 AlaArgLysAlaThrValMetLysIleProValProThrThrProArgGlnValArgGlu 419
QY 3420 TTTTGGGGACAGCTGGATTTCGAGACTGTGATCCCGGGTTTGGACTTTAGCAGCC 3479
Db 420 PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaSerLeuAlaAla 439
QY 3480 CCACCTACCCGCTAACCAAGAAAAGGGAAATTCCTCGGGCTCTGAGCACCAGAAG 3539
Db 440 ProLeuTyrProLeuThrLysGluSerIleProPheIleTrpThrGluGluHisGlnGln 459
QY 3540 GCATTTGATGCTATCAAAAGGCCCTCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTA 3599
Db 460 AlaPheAspHisIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspLeu 479
QY 3600 ACTAAACCCCTTACCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGGGAGTTTAAAC 3659
Db 480 ThrLysProPheThrLeuTyrIleAspGluArgAlaGlyValAlaArgGlyValLeuThr 499
QY 3660 CAAACCTTAGGACCATGGAGAAGACCTGTGCTTACCTGTCAAGAAAGCTCGATCCTGTA 3719
Db 500 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProVal 519
QY 3720 GCCAGTGGTGGCCCATATGCCATGAGGCTATCGAGCTGTGGCCATATCTGTGCAAGGAC 3779
Db 520 AlaSerGlyTrpProThrCysLeuLysAlaValAlaAlaValAlaLeuLeuLysAsp 539
QY 3780 GCTGACAAATGACTTTGGACAGAATATAACTGTATAGCCCTCCATCGATTTGGAGAAC 3839
Db 540 AlaAspLysLeuThrLeuGlyGlnAsnValThrValIleAlaSerHisSerLeuGluSer* 559
QY 3840 ATGTTTGGCGAGCCCGACCATGGATGACCAACGCGCGCATGACCCACTATCAAGC 3899
Db 560 IleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSer 579
QY 3900 CTGCTTCTCACAGAGGGTTCAGTTTCGCTCCACCGCGCTCTCAACCTCGCACTCTT 3959

Db 580 LeuLeuLeuAsnGluArgValSerPheAlaProProAlaValLeuAsnProAlaThrLeu 599
QY 3960 CTCCTCGAAGAGACTGATGAACAGTACCTCATGATTGCCATCACTATTGATTGAGGAG 4019
Db 600 LeuProValGluSerGluAlaThrProValHisArgCysSerGluIleLeuAlaGluGlu 619
QY 4020 ACTGGGGTCCGAAGGACCTTACAGACATACCCCTGACTCGAGAAAGTCTTAACCTGGTTC 4079
Db 620 ThrGlyThrArgArgAspLeuGluAspGlnProLeuProGly--ValProThrTrpTyr 638
QY 4080 ACTGACCGAGACGACTATGTGGTGAAGGTAAAGAGATGCTGGGGCGGGTGGTGCAGC 4139
Db 639 ThrAspGlySerSerPheIleThrGluGlyLysArgAlaGlyAlaProIleValAsp 658
QY 4140 GGCACCGCAGCATCTGGGCCAGCAGCCTCCCGAAGGAATCTTCAGCACAAAAGCTGAG 4199
Db 659 GlyLysArgThrValTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaGlu 678
QY 4200 CTCATGGCCCTCACCAAGCTTTGCGGCTGGCCGAAGGGAATTCATATAACATTTATACG 4259
Db 679 LeuValAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysAsnIleAsnIleTyrThr 698
QY 4260 GACAGCAGGTATGCCCTTTCGCACTGCACAGTACATGGGGCCATCTATATAACAAAGGGG 4319
Db 699 AspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyAlaIleTyrLysGlnArgGly 718
QY 4320 TTGCTTACCTCAGCAGGAGGAAATAAAGAACAAAGAGGAATTCCTAAGCCCTATTAGAA 4379
Db 719 LeuLeuThrSerAlaGlyLysAspIleLysAsnLysGluGluIleLeuAlaLeuLeuGlu 738
QY 4380 GCGTACATTTTACCAAAAGGCTATTATATACCTGTCTCGACATCAGAAAGCTATAA 4439
Db 739 AlaIleHisLeuProArgValAlaIleIleHisCysProGlyHisGlnArgGlySer 758
QY 4440 GATCTCATATCCAGAGAAACCGATGGCTGACCGGGTTCGCAAGCAGGAGCCAGCGT 4499
Db 759 AsnProValAlaThrGlyAsnArgAlaAspGluAlaAlaLysGlnAlaLeuSer 778
QY 4500 GTTAACTTCTGCTATAATAGAAATGCCCAAGCCCA--CAACCCAGCACACAGTAC 4556
Db 779 ThrArgValLeuAlaGlyThrThrLysProGlnGluProIleGluProAlaGlnGluLys 798
QY 4557 ACCTTAGAAGACTGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAAGG 4616
Db 799 ThrArgPro----- 801
QY 4617 ACCTGCTATACCTCAGATCGGAAGAAATCTCTGCCCAACAAAGAGGTTAGAAATATGC 4676
Db 802 -----ArgGluLeuThrProAspArg--GlyLysGluPheIle 813
QY 4677 CAACGATACATCGCTAACCCACCTAGGAACATAAACACCTGACGAGCTTTGGTCAAGAACA 4736
Db 814 LysArgLeuHisGlnLeuThrHisLeuGlyProGluLysLeuLeuGlnLeuValAsnArg 833
QY 4737 TCCCTTATCATGTTCTGAGGCTACCAGGAGTGGCTGACTCGGTGCTCAACATTTGTGTG 4796
Db 834 ThrSerLeuLeuIleProAsnLeuGlnSerAlaValArgGluValThrSerGlnCysGln 853
QY 4797 CCGTGCACCTGGTTAATCTAATCTCTCCAGAAATCCCTCCAGGGAAGAGACATAAGGGA 4856
Db 854 AlaCysAlaMetThrAsnAlaValThrThrTyrArgGluThrGlyLysArgGlnArgGly 873
QY 4857 AGCCACCCAGGCGCTCACCTGGGAAGTGGACTTCACCTGAGTAAAGCCGCTAATAACGGA 4916
Db 874 AspArgProGlyValTyrTrpGluValAspPheThrGluIleLysProGlyArgTyrGly 893
QY 4917 AACAAATACCTATTGGTTTTGTAGACACCTTTTTCAGGATGGGTAGAGGCTTATCCTACT 4976
Db 894 AsnLysTyrLeuLeuValPheIleAspThrPheSerGlyTrpValGluAlaPheProThr 913
QY 4977 AAGAAAGAGACTTCAACCTGGTGGCTAAATAAATACTCGAAGAAATTTTCCAGATTT 5036
Db 914 LysThrGluThrAlaLeuIleValCysLysIleLeuGluGluIleLeuProArgPhe 933

```
QY 5037 GGAATACCTAAGGTAATAGGTCAGACAAATGGTCCAGCTTTTGTGGCCCGAGTAAGTCAG 5096
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 934 GlyIleProLysValLeuGlySerAspAsnGlyProAlaPheValAlaGlnValSerGln 953
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5097 GGACTGGCCCAAGATATTGGGGATTGATTGGAACTGTCATTGTGCATACAGACCCCAAGC 5156
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 954 GlyLeuAlaThrGlnLeuGlyIleAsnTrpLysLeuHisCysAlaIleArgProGlnSer 973
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5157 TCAGCAGAGGTAGAGAGGATGAATAGAACCATTAAGAGACCCCTTACTAAATTGACCGCG 5216
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 974 SerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuAlaLeu 993
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5217 GAGACTGGCTTAATGATTGATAGCTCTCTGCGCTTTGTGCTTTTGTAGGTTAGAAC 5276
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 994 GluThrGlyLysAspTrpValThrLeuLeuProLeuAlaLeuLeuAlaArgAsn 1013
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5277 ACCCTCGACAGTTTGGGCTGACCCCTATGAATTACTCTACGGGGGACCCGCCCATTTG 5336
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1014 ThrProGlyArgPheGlyLeuThrProTyrGluIleLeuTyrGlyGlyProProProIle 1033
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5337 GTAGAAATTGCTTCTGTACATAGTGTGACGTGCTGCTTTCCAGCCCTTTGTTCTCTAGG 5396
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1034 LeuGluSerGlyGluThrLeuGlyProAspAspArgPheLeuProValLeuPheThrHis 1053
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5397 CTCAGGACCTTGTAGTGGGTGAGCAACAGCGGTGGAGGCACTCCGGGAGGCGCTACTCA 5456
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1054 LeuLysAlaLeuGluIleValArgThrGlnIleTrpAspGlnIleLysGluValTyrLys 1073
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5457 GGAGGAGAGACTTGCAGATCCCATCGTTTCCAAAGTGGGAGATTCAGTCTACGTTAGA 5516
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1074 ---ProGlyThrValThrIleProHisProPheGlnValGlyAspGlnValLeuValArg 1092
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5517 CGCCACCGTCAGGAAACCTTCAGACTCGGTGGGAAGGCGCCTTATCTCGTACTTTTGACC 5576
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1093 ArgHisArgProSerSerLeuGluProArgTrpLysGlyProTyrLeuValLeuLeuThr 1112
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5577 ACACCAAGCGCTGGAAGTGAAGGAAATCTCCACCTGGATCCATGATCCAGCTTAA 5636
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1113 ThrProThrAlaValLysValAspGlyIleAlaIleValAlaIleValHisAlaSerHisLeuLys 1132
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5637 CCGCGCGCACCT-----CCGATTTCGGGTGGGAAGCCGAAAGAGATGAAATCCCTT 5690
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1133 ProAlaProProSerAlaProAspGluSerTrpGluLeuGluLysThrAspHisProLeu 1152
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 5691 AAGCTTCGCTCCATCGC 5708
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1153 LysLeuArgIleArgArg 1158
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 3
GNMVG
HIV-1 retropepsin (EC 3.4.23.16) - AKV murine leukemia virus
N:Contains: nuclelease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polymera
C:Species: AKV murine leukemia virus
A:Note: host Mus spp. (mouse)
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 31-Dec-2004
C:Accession: B92995; B93448; A03957
R:Herr, W.
J. Virol. 49, 471-478, 1984
A:Title: Nucleotide sequence of AKV murine leukemia virus.
A:Reference number: A92995; MUID:84115072; PMID:6319746
A:Accession: B92995
A:Molecule type: genomic RNA
A:Residues: 1-1196 <HER>
A:Cross-references: UNIPROT:Q9E7M1; UNIPROT:Q83362; UNIPROT:Q9J8E2; UNIPROT:Q90RL4; UNIR
R:Herr, W.; Corbin, V.; Gilbert, W.
Nucleic Acids Res. 10, 6931-6944, 1982
A:Title: Nucleotide sequence of the 3' half of AKV.
A:Reference number: A93448; MUID:83090450; PMID:6294621
A:Accession: B93448
A:Molecule type: DNA
A:Residues: 353-995, 997-1196 <HE2>
A:Cross-references: UNIPARC:UPI0000131F01
```

C;Comment: This protein is synthesized as a gag-pol polyprotein.
C;Comment: The pol polyprotein contains reverse transcriptase and possibly a nuclease or
C;Genetics:
A;Gene: pol
C;Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polyprotein; reverse
F;3-102/Product: retropepsin #status predicted <RTP>
F;27/Active site: Asp (shared with dimeric partner) #status predicted

Alignment Scores:

Pred. No.:	4,09e-238	Length:	1196
Score:	3999.00	Matches:	753
Percent Similarity:	77.2%	Conservative:	172
Best Local Similarity:	62.9%	Mismatches:	259
Query Match:	27.3%	Indels:	14
DB:	1	Gaps:	9

US-10-723-552-3 (1-8132) x GNMVG (1-1196)

```
QY 2160 GGGAGACGGGTTTCGGACCCCTCCCGAGCCCGGGTAACTTTGAAGTGGAGGGCAA 2219
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1 GlyGlyGlnGlyGlnGluProProGluProArgIleThrLeuThrValGlyGln 20
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2220 CCAGTTGAGTTCTCTGGTTGATACCGGAGCGCAACATTCACTGCTACTACAGCCATTAGGA 2279
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 21 ProValThrPheLeuValAspThrGlyAlaGlnHisSerValLeuThrGlnAsnProGly 40
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2280 AAATAAGATAAAAATCTCGGTGATGGTGCCACAGGGCAACAAACAGTATCCATGG 2339
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 41 ProLeuSerAspArgSerAlaTrpValGlnGlyAlaThrGlyGlyLysArgTyrArgTrp 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2340 ACTACCCGAAGACAGTTGACTTGGGAGTGGGAGTGGCAACCACTGTTTCTGGTCATA 2399
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 ThrThrAspArgLysValHisLeuAlaThrGlyLysValThrHisSerPheLeuHisVal 80
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2400 CCTGAGTCCAGACACCCCTCTTAGGTAGACACTTATTGACCAAGATGGGAGCACAAATT 2459
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 81 ProAspCysProTyrProLeuLeuGlyArgAspLeuLeuThrLysLeuLysAlaGlnIle 100
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2460 TCTTTTGA---CAAGGGAACAGAAAGTCTCTGCMAATTAACAACTATCCTGTTG 2516
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 101 HisPheGluGlySerGlyAlaGlnValGlyProLysGlyGlnProLeuGlnValLeu 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2517 ACCTCCAAATAGATGACGAATATCGACTATCTCTCCCTAGTAAAGCCTGATCAAAAT 2576
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 ThrLeuAsnLeuGluAspGluTyrArgLeuTyrGluThrSerAlaGluProGluValSer 140
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2577 ATA---CAATTCCTGTTGGAACAGTTTCCCAAGCCTGGCGAGAAACCGCAGGATGGGT 2633
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 141 ProGlySerThrTrpLeuSerAspPheProGlnAlaTrpAlaGluThrGlyGlyMetGly 160
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2634 TTGGCAAGCAAGTTCCCCACAAAGTTATTCACTGAAGCCAGTGCACACCAAGTGTCA 2693
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 161 LeuAlaValArgGlnAlaProLeuIleProLeuLysAlaThrSerThrProValSer 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2694 GTCAGACAGTACCCTTTGAGTAAAGAGCTCAAGAAGAAATTCGGCCGCATGTCCAAAGA 2753
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 181 IleLysGlnTyrProMetSerGlnGluAlaLysLeuGlyLysProHisIleGlnArg 200
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2754 TTAATCCAACAGGCGCATCTCTAGTTCTCTGTCCTCAATCTCCCTGGAACTATCCCTGCTACCG 2813
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 201 LeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeuLeuPro 220
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2814 GTTAGAAAGCCTGGGACTAATGACTATCGACCACTACAGACTTCAGAGAGGTCAATAAA 2873
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 221 ValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLys 240
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2874 CGGTGCGAGGATATACACCAACAGTCCCGAACCTTATTAACCTTGTGTGCTCTCCCA 2933
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 241 ArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeuPro 260
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 2934 CCCCACCGAGCTGGTATACAGTATTGAGCTTAAGAGATCCCTTTCTGCTCGAGATTA 2993
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 261 ProSerHisArgTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeu 280
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

QY 2994 CACCCCTAGCCAAACCACTTTTGGCTTCCGATCGAATGAGAGATCCAGGTACGGGAAGAACC 3053
DB HisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyMetGlyIleSer 300
QY 3054 GGGCAGCTCACCTCGACCGGACCTGCCCAAGGTTCAAGACTCCCGACCATCTTTGAC 3113
DB GlyGlnLeuThrTrpThrArgLeuProGlnGlyPheIysAsnSerProThrLeuPheAsp 320
QY 3114 GAAGCCCTACACAGAGACCTGGCCAACTTCAGAGATCCAAACCCCTCAGGTGACCCCTCCTC 3173
DB GluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuIleLeuLeu 340
QY 3174 CAGTACCTGATGACCTGCTTCTGGCGGAGCCACCAACAGGACTCTCTAGAGGACG 3233
DB GlnTyrValAspAspIleLeuLeuAlaAlaThrSerGluLeuAspCysGlnGlnGlyThr 360
QY 3234 AAGGCCTACTGCTGGAATGCTGACCTAGGCTACAGAGCCTCTGCTAGAGGCCCGCAG 3293
DB ArgAlaLeuLeuLeuThrLeuGlyAsnLeuGlyTyrArgAlaSerAlaIleIleLeuAlaGln 380
QY 3294 ATTTGCAGGAGAGAGGTAAACATCTCTGGGGTACAGTTTGGGGACGGGACGATGGCTG 3353
DB LeuCysGlnIysGlnValIysTyrLeuGlyTyrLeuLeuIysGluGlyGlnArgTrpLeu 400
QY 3354 ACGGAGCGACGGAAGAAACTGTAGTCAGATACCGGCCCCCAACACGACCAACAAATG 3413
DB ThrGluAlaArgIysGluThrValMetGlyGlnProThrProIysThrProArgGlnLeu 420
QY 3414 AGAGATTTTGGGACAGCTGGATTTTGGAGCTGTGATCCCGGGTTTGGACCTTA 3473
DB ArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGluMet 440
QY 3474 GCAGCCCACTTACCCGCTAAACCAAGAAAAGGGGAATTCCTGGGCTCCTGAGCAC 3533
DB AlaAlaProLeuThrProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAspGln 460
QY 3534 CAGAAGCATTTGATGCTATCAAAAAGGCCCTCTGAGCGCACCTGCTCTGGCCCTCCCT 3593
DB GlnLysAlaTyrGlnGluIleIysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeuPro 480
QY 3594 GAGTAACTAAACCTTTACCTTTATGTCGATGAGGTAAAGGATAGCCCGGGAGTT 3653
DB AspLeuThrIysProPheGluLeuPheValAspGluIysGlnGlyTyrAlaIysGlyVal 500
QY 3654 TTAACCCAAACCTTAGGACCATGAGAACCTGTGCTACCTGTCAAGAGAGCTCGAT 3713
DB LeuThrGlnIysLeuGlyProThrArgArgProValAlaTyrLeuSerIysIysLeuAsp 520
QY 3714 CTTGTAGCCAGTGTGGCCCATATGCTGAAGGCTATCGCAGCTGTGGCCATCTGGTC 3773
DB ProValAlaAlaGlyTrpProProCysLeuArgMetValAlaAlaIleAlaValLeuThr 540
QY 3774 AAGGACGCTGACAAATGACTTTGGGACAGAAATATTAAGTAAAGCCCCCATGATG 3833
DB LysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAlaVal 560
QY 3834 GAGAACATCTCTCGGAGCCCGCAGACCGATGATGACCAACCGCCGATGACCCACTAT 3893
DB GluAlaLeuValIysGlnProProAspArgTrpLeuSerAsnAlaArgMetThrHisTyr 580
QY 3894 CAAAGCCTGCTCTC---ACAGAGAGGGTCACGCTCTCGCTCCACGCGCTCTCAACCCCT 3950
DB GlnAlaMetLeuLeuAspThrAspArgValGlnPheGlyProValValAlaLeuAsnPro 600
QY 3951 GCCACTCTTCTGCTGAAGAGACTGATGAACACGATGACTCATGATGCCATCACTATTG 4010
DB AlaThrLeuLeuPro---LeuProGluGluGlyAlaProHisAspCysLeuGluIleLeu 619
QY 4011 ATTGAGGACTGGGCTCCGAGGACCTTACAGACATACCGCTGACTGGAGAGTGCTA 4070
DB AlaGluThrHisGlyThrArgProAspLeuThrAspGlnProIleProAspAlaAspHis 639

QY 4071 ACCTGGTTCTACTGACGGAACGACGCTATGTGTGGAAGTAAGAGATGGCTGGGGCGCGC 4130
DB ThrTrpTyrThrAspGlySerSerPheLeuGlnGlyGlnArgLysAlaGlyAlaAla 659
QY 4131 GTGGTGAACGGGACCGCAGACTCTGGCCAGACGCTCTGCCGAAGGAACCTTCAGCAAA 4190
DB ValThrThrGluThrGluValIleTrpAlaArgAlaLeuProAlaGlyThrSerAlaGln 679
QY 4191 AAGGCTCAGCTCATGGCCCTCACCAAGCTTTGGGCTGGCCGAAGGGAATTCATATAAC 4250
DB ArgAlaGluLeuIleAlaLeuThrGlnAlaLeuIysMetAlaGluGlyLysArgLeuAsn 699
QY 4251 ATTATATAGACACGACGATGCTTTGCGACTGCACACGTACATGATGGGCCATCTATAAA 4310
DB ValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyGluIleTyrArg 719
QY 4311 CAAAGGGGTTGCTTACCTCAGCAGGAGGGAATTAAGAACAAAGAGGAATTCCTAAGC 4370
DB ArgArgGlyLeuLeuThrSerGluGlyArgGluIleLysAsnLysSerGluIleLeuAla 739
QY 4371 CTATTAGAAGCCGTACATTTACCAAAAGCTAGCTATTATATACACTGCTCTGGACATCAG 4430
DB LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysLeuGlyHisGln 759
QY 4431 AAAGCTAAAGATCTCATATCCAGAGGAACACGATGCTGACCGGTTGCCAAGCAGCCA 4490
DB LysGlyAspSerAlaGluAlaArgGlyAsnArgLeuAlaAspGlnAlaAlaArgGluAla 779
QY 4491 GCC-----CAGGGTGTAACTTCTGCTTATATAAGAAATGCCCAAGCCCA 4538
DB AlaIleLysThrProProAspThrSerThrLeuLeuIleGluAspSerThrProTyrThr 799
QY 4539 GAACCCAGACGACAGTACACCTTAGAAGACTGCGCAGAGATATAAAGATAGACCACTTC 4598
DB ProAlaTyrPheHisTyrThrGluThrAspLeuLysLeuArgGluLeuGlyAlaThr 819
QY 4599 TCTGAGACTCCGGAAGGAGCCTGCTATACCTCAGATGGGAAGAAATCTCGCCCAAAA 4658
DB TyrAsnGlnSerIysGly--TyrTrpValPheGlnGlyLysProValMetProAspGln 838
QY 4659 GAAGGTTAGATATATGTCMAACAGATACATGCTTAACCCACCTAGGAACCTAACACCTG 4718
DB PheValPheGluLeuLeuAspSerLeuHisArgLeuThrHisLeuGlyTyrGlnLysMet 858
QY 4719 CAGCAGTTGTC-----AGAACATCCCTTATCATGTTCTGAGGCTACCAAGAGTGCT 4772
DB LysAlaLeuLeuAspArgGlyGluSerProTyrTyrMetLeuAsnArgAspLysThrLeu 878
QY 4773 GACTCGGTGCTCAAAACATTTGTGTCCTGCTGCTGCTTAATCTTCCAGAAATG 4832
DB GlnTyrValAlaAspSerCysThrValCysAlaGlnValAsnAlaSerLysAlaLysIle 898
QY 4833 CCTCCAGGAGAGACTAAGGGAAGCCACCCAGGCGCTCAGTGGGAAGTGGACTTCCT 4892
DB GlyAlaGlyValArgValArgGlyHisArgProGlySerHisTrpGluIleAspPheThr 918
QY 4893 GAGGTAACCCGCTAAATACGGAACAAATACCTATTGTTTCTGAGACACTTTCA 4952
DB GluValLysProGlyLeuTyrGlyTyrLysTyrLeuLeuValPheValAspThrPheSer 938
QY 4953 GGATGGGTAGAGCTTATCTACTAGAAAGAGACTTCAACCGCTGGCTGGCTAAAAAATA 5012
DB GlyTrpValGluAlaPheProThrLysArgGluThrAlaArgValValSerLysLysLeu 958
QY 5013 CTGGAAGAAATTTTCCAAAGATTTGGAATACCTAAGTAAATAGGTCAGACAATGGTCCA 5072
DB LeuGluGluIlePheProArgPheGlyMetProGlnValLeuGlySerAspAsnGlyPro 978
QY 5073 GCTTTTGTGTCAGTAACTAGGACTGGCCAGATATTGGGATTCATTGGAACCTG 5132
DB AlaPheThrSerGlnValSerGlnSerValAlaAspLeuLeuGlyIleAspTrpLysLeu 998
QY 5133 CATTTGTGCATACAGACCCCAAGCTCAGGACAGGTAGACAGGATGAATAGAACCTATAA 5192

```

Db      999  HisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLys 1018
Qy      5193  GAGACCTTACTAAATGACCGCGGAGACTGGCGTTAATGATGGATAGCTCTCTGCCCC 5252
Db      1019  GluThrLeuThrLysLeuThrLeuAlaAlaGlyThrArgAspTrpValLeuLeuLeuPro 1038
Qy      5253  TTTGTGCTTTTAGGGTTAGGAACACCCCTGGACAGTTTGGCGTACCCCTATGAATTA 5312
Db      1039  LeuAlaLeuTyrArgAlaArgAsnThrProGlyProHisGlyLeuThrProTyrGluile 1058
Qy      5313  CTCTACGGGGGACCCCGCCCTAGTGGTAAATTCCTTCTGTACATAGTCTGACGCTG 5372
Db      1059  LeuTyrGlyAlaProProLeuValAsnPheHisAspProAspMetSerGluLeuThr 1078
Qy      5373  CTTTCCAGCCTTGTCTCTAGGCTCAAGGCACTTGAAGTGGGTGAGACAAGAGCGTGG 5432
Db      1079  AsnSerProSerLeuGlnAlaHisLeuGlnAlaLeuGlnThrValGlnArgGluIleTrp 1098
Qy      5433  AGGCAACTCCGGGAGGCTACTCAGGAGGAGGAGACTTGCAG---ATCCACATCGTTTC 5489
Db      1099  LysProLeuAlaGluAlaTyrArgAspGlnLeuAspGlnProValIleProHisProPhe 1118
Qy      5490  CAAGTGGGAGATTCACTCTAGCTTAGAGCCGCCCGTGCAGGAAACCTCGAGACTCGGTGG 5549
Db      1119  ArgIleGlyAspSerValTrpValArgArgHisGlnThrLysAsnLeuGluProArgTrp 1138
Qy      5550  AAGGCGCCTTATCTGCTATCTTTCACACACCAACGCTGTGAAAGTCGAAGGAATCTCC 5609
Db      1139  LysGlyProTyrThrValLeuLeuThrThrProThrAlaLeuLysValAspGlyIleSer 1158
Qy      5610  ACCTGGATCCATGATCCAGTTAAACCGCG-----CCACCTCCCGATTCCGGGTGG 5663
Db      1159  AlaTrpIleHisAlaAlaHisValLysAlaAlaThrThrProProIleLysProSerTrp 1178
Qy      5664  AAAGCCCAAGACGTAAATCCCTTAAGCTTCGCTCCATCCGCTCGGTTCCT 5717
Db      1179  ArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGlyAlaPro 1196

RESULT 4
S35475
pol polyprotein - Friend murine leukemia virus (strain PVC-211)
N;Contains: nuclease; proteinase; reverse transcriptase
C;Species: Friend murine leukemia virus
A;Variety: strain PVC-211
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S35475
R;Remington, M.P.; Hoffman, P.M.; Ruscetti, S.K.; Masuda, M.
Nucleic Acids Res. 20, 3249, 1992
A;Title: Complete nucleotide sequence of a neuropathogenic variant of Friend murine leuk
A;Reference number: S35474; MUID:92319660; PMID:1620621
A;Accession: S35475
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-1204 <REM>
A;Cross-references: UNIPROT:P26808; UNIPARC:UPI0000131F06; EMBL:M93134; NID:G331898; PII
A;Experimental source: strain PVC-211
A;Note: the authors translated the stop codon at position 5 as Gln
C;Genetics:
C;Superfamily: pol polyprotein
C;Keywords: polyprotein; reverse transcriptase

Alignment Scores:
Pred. No.: 1,03e-237 Length: 1204
Score: 3992.50 Matches: 757
Percent Similarity: 76.1% Conservative: 162
Best Local Similarity: 62.7% Mismatches: 257
Query Match: 27.3% Indels: 31
DB: 2 Gaps: 11

US-10-723-552-3 (1-8132) x S35475 (1-1204)

```

```

Qy      2154  GATTAGGGAGACGGGTTTCGGACCCCTCCCGAGCCAGCGTAACCTTTGAAGTGGAG 2213
Db      4      AspGlnGlyGlyGlnGlyGlnProProGluProArgIleThrLeuLysValGly 23
Qy      2214  GGCAACCACTGAGTTCTCGTTGATACCGGAGCGAAACATTCAGTGTACTACACCA 2273
Db      24      GlyGlnProValThrPheLeuValAspThrGlyAlaGlnHisSerValLeuThrGlnAsn 43
Qy      2274  TTAGGAAACTAAAGATAAAATCCTGGTGTATGGTGGCCACAGGCAACACAGTAT 2333
Db      44      ProGlyProLeuSerAspLysSerAlaTrpValGlnGlyAlaThrGlyGlyLysArgTyr 63
Qy      2334  CCATGGACTACCCGAACACAGTTGACTCGGAGTGGGACGGTGAACCCACTCTGTTCTG 2393
Db      64      ArgTrpThrThrAspArgValHisLeuAlaThrGlyLysValThrHisSerPheLeu 83
Qy      2394  GTCATACCTGAGTCCCGACACCCCTCTTATAGGTAGACATTATTGACCAAGATGGGAGCA 2453
Db      84      HisValProAspCysProTyrProLeuLeuGlyArgAspLeuLeuThrLysLeuLysAla 103
Qy      2454  CAAATTTCTTTTGA---CAAGGAAACACAGAGTGTCTCAATAATAACAACCTATCACT 2510
Db      104      GlnIleHisPheGluGlySerGlyAlaGlnValValGlyProMetGlyGlnProLeuGln 123
Qy      2511  GTCTTGACCCCTCCAATTAGATGACGAATATCGACTATACTCTCCCTAGTAAAGCCCTGAT 2570
Db      124      ValLeuThrLeuAsnIleGluAspGluTyrArgLeuHisGluThrSerLysGlyProAsp 143
Qy      2571  CAAAATATA---CAATTTCTGTTGGAACAGTTTCCCCAACCTGGGAGAAACCGCAGGG 2627
Db      144      ValProLeuGlySerThrTrpLeuSerAspPheProGlnAlaTrpAlaGluThrGlyGly 163
Qy      2628  ATGGTTTGGCAAGCAAGTTTCCCCCAAGTTATTCACTGAAGCCAGTGCACACCA 2687
Db      164      MetGlyLeuAlaValArgGlnAlaProLeuIleIleProLeuArgAlaIleSerThrPro 183
Qy      2688  GTGTCTGACGACAGTACCCCTTGAGTAAAGAGCTCAAGAGAAATTCGCGCCCATGTC 2747
Db      184      ValSerIleLysGlnTyrProMetSerArgGluAlaArgLeuGlyIleLysProHisIle 203
Qy      2748  CAAAGATTAAATCCAACAGGCGATCTTAGTTCTCTGTCCAATCTCCCTGGAATCTCCCTG 2807
Db      204      GlnArgLeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeu 223
Qy      2808  CTACCGTTAGAAAGCCTCGGACTAATGACTATCGACCACTACAGGCTTGAAGAGGTC 2867
Db      224      LeuProValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluVal 243
Qy      2868  AATAAAGCGGTGCAGGATATACACCCCAAGTCCCGAACCTTATAACCTCTTGTGTGCT 2927
Db      244      AsnLysArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGly 263
Qy      2928  CTCCACCCCAACGGAGCTGTATACAGTATTGAGATTAAAGGATGCTTCTTCTGCTG 2987
Db      264      LeuProProSerHisGlnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeu 283
Qy      2988  AGATTACACCCCTACGCCAACCACTTTTTCCTTCAATGGAGAGATCCAGGTACGGGA 3047
Db      284      ArgLeuHisProThrSerGlnSerLeuPheAlaPheGluTrpArgAspProGluMetGly 303
Qy      3048  AGAACCGGCGAGCTCACTCGACCCGACTCCCAAGGTTCAAGAACTCCCCACCATC 3107
Db      304      IleSerGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeu 323
Qy      3108  TTTGACCAAGCCCTACACAGAGACTCGGCCCACTTCAGGATCCAAACCCCTCAGGTGACC 3167
Db      324      PheAspGluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuLe 343
Qy      3168  CTCCTCAGTACGTGGATGACCTGCTTCTGCGGGAGGCCACCAACAGGACTGTTAGAA 3227
Db      344      LeuLeuGlnTyrValAspAspLeuLeuAlaAlaThrSerGluLeuAspCysGlnGln 363

```


Db 1075 ProAspProAspMetAlaLysValThrHisAsnProSerLeuGlnAlaHisLeuGlnAla 1094

Qy 5406 CTTGAGTGGGTGAGACAAACGAGCGTGGAGCAACTCCGGGAGGCCTACTCAGGAGGAGGA 5465
||| |||::: ||||| ||| |||||

Db 1095 LeuTyrLeuValGlnHisGluValTrpArgProLeuAlaAlaTyrGlnGluGlnLeu 1114
||| |||::: ||||| ||| |||||

Qy 5466 GACTTGCAG---ATCCACATCGTTTCCAAAGTGGAGATTCACTGCTACGTTAGACGCCAC 5522
||| |||::: ||||| ||| |||||

Db 1115 AspArgProValValProHisProPheArgValGlyAspThrValTrpValArgArgHis 1134
||| |||::: ||||| ||| |||||

Qy 5523 CGTGCAGGAACCTCGAGACATCGGTGGAAGGGCCCTTATCTCGTACTTTTGGACACACCA 5582
:: ||| |||::: ||||| ||| |||||

Db 1135 GlnThrLysAsnLeuGluProArgTyrGlyProTyrThrValLeuLeuThrThrPro 1154
||| |||::: ||||| ||| |||||

Qy 5583 ACGGCTGTGAAAGTCGAAGGATCTCCACCTGGATCCATCCATCCACGTTAAACCGGCG 5642
||| |||::: ||||| ||| |||||

Db 1155 ThrAlaLeuLysValAspGlyIleAlaAlaTrpIleHisAlaAlaHisValLysAlaAla 1174
||| |||::: ||||| ||| |||||

Qy 5643 -----CCACCTCCCGATTCCGGGTGGAAAGCCGAAAGACTGAAAAATCCC 5687
||| |||::: ||||| ||| |||||

Db 1175 AspThrLysIleGluProProSerGluSerThrTrpArgValGlnArgSerGlnAsnPro 1194
||| |||::: ||||| ||| |||||

Qy 5688 CTTAAGCTTCGCTCCATCGC 5708
||| |||::: ||||| ||| |||||

Db 1195 LeuLysIleArgLeuThrArg 1201
||| |||::: ||||| ||| |||||

RESULT 5

S70393
pol polyprotein - Friend murine leukemia virus (strain FB29)
N:Contains: nuclease; proteinase; reverse transcriptase
C:Species: Friend murine leukemia virus
A:Variety: strain FB29
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
R/Perryman, S.; Nishio, J.; Chesebro, B.
C/Accession: S70393
Nucleic Acids Res. 19, 6950, 1991
A/Title: Complete nucleotide sequence of Friend murine leukemia virus, strain FB29.
A/Reference number: S70393; MUID:92107687; PMID:1762923
A/Accession: S70393
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: genomic RNA
A/Residues: 1-1204 <PER>
A/Cross-references: UNIPROT:P26809; UNIPARC:UPI0000131F05; EMBL:Z11128; NID:961547; PDB:
A/Experimental source: strain FB29
A/Note: the authors translated the stop codon at position 5 as Gln
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C/Genetics:
A/Gene: pol
C/Superfamily: pol polyprotein
C/Keywords: polyprotein; reverse transcriptase

[illegible]

Db	181	IleLysGlnTyrProMetSerGlnGluAlaLysLeuGlyIleLysProHisIleGlnArg	200
Qy	2754	TTAATCAACAGCGCATCTCTAGTTCTCTCCAATCTCCCTGGAATATCTCCCTGCTACCG	2813
Db	201	LeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeuLeuPro	220
Qy	2814	GTTAGAAGCCTGGGACTAATAGTATPCGACAGTACAGGACTTGGAGAGGTCAATAAA	2873
Db	221	ValLysLysProGlyThrAsnAspTyrArgProValGlnGlyLeuArgGluValAsnLys	240
Qy	2874	CGGGTCAGGATATACACCCAAACAGTCCCGAACCTTATACCTCTCTGTGTCTCTCCCA	2933
Db	241	ArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeuPro	260
Qy	2934	CCCCAAGCGAGCTGGTATACATCAGTATTGGACTTAAAGGATCGCTTCTCTGCTGAGATTA	2993
Db	261	ThrSerHisArgTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeu	280
Qy	2994	CACCCCACTAGCCAAACACATTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACC	3053
Db	281	HisProThrSerGlnProLeuPheAlaSerGluTrpArgAspProGlyMetGlyIleSer	300
Qy	3054	GGCGAGCTACCTGGACCCGACCTGCCCGAGGGTTCAGAACTCCCGCGACCATCTTTGAC	3113
Db	301	GlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAsp	320
Qy	3114	GAAGCCCTACACAGAGACTGGCCNACTTCAGGATCCAAACCCCTCAGGTGACCCCTCCTC	3173
Db	321	GluAlaLeuHisArgGlyLeuAlaAspPheArgIleGlnHisProAspLeuIleLeuLeu	340
Qy	3174	CAGTACGTGGATGACCTGCTTCTGGCGGGAGCCACCAACAGGACTCTCTTAGAAGGCACG	3233
Db	341	GlnTyrValAspAspLeuLeuLeuAlaAlaThrSerGluLeuAspCysGlnGlnGlyThr	360
Qy	3234	AAGCACTACTGCTGGAATTTGCTGCTACCTAGGCTACAGAGCCTCTGCTAAGAGGCCACG	3293
Db	361	ArgAlaLeuLeuLysThrLeuGlyAsnLeuGlyTyrArgAlaSerAlaLysLysAlaGln	380
Qy	3294	ATTTCGAGGAGAGAGGTAAACATACTCTGGGTACAGATTTTCGGGACGGCAGCGATGGCTG	3353
Db	381	IleCysGlnLysGlnValLysTyrLeuGlyTyrLeuLeuArgGluGlyGlnArgTrpLeu	400
Qy	3354	ACGAGGACCGGAAGAAAATCTAGTCAGATACCGGCCCCCAACACAGGCCAAACAAATG	3413
Db	401	ThrGluAlaArgLysGluThrValMetGlyGlnProThrProLysThrProArgGlnLeu	420
Qy	3414	AGAGAGTTTTTGGGACAGCTGGAATTTTGACAGACTGTGGATCCCGGGTTTGGACCTTA	3473
Db	421	ArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProArgPheAlaGluMet	440
Qy	3474	GCAGCCCACTCTACCCGCTAACCAAGAAAAAGGGGAATTTCTCTGGGCTCTCTGAGCAC	3533
Db	441	AlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAspGln	460
Qy	3534	CAGAAGCATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCT	3593
Db	461	GlnLysAlaTyrHisGluIleLysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeuPro	480
Qy	3594	GAGTAACTAAACCTTTTACCTTTATGTGGATGAGCGTAAAGGAGTAGCCCGGGGAGTT	3653
Db	481	AspLeuThrLysProPheGluLeuPheValAspGlnLysGlnGlyTyrAlaLysGlyVal	500
Qy	3654	TTAAACCCAAACCTTAGGACCATGGAGAAGACCTGTGCGCTACCTGTCAAAAGAAGCTCGAT	3713
Db	501	LeuThrGlnLysLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAsp	520
Qy	3714	CCTGTAGCCAGTGGTTGGCCCATATGCTCGAAGGCTATCGCAGCTGTGGCCATACTGGTC	3773
Db	521	ProValAlaAlaGlyTrpProCysLeuArgMetValAlaAlaIleAlaValLeuThr	540
Qy	3774	AAGGACCTGACAAATTCATCTTTGGGACAGAAATAAATACTGTAATAGCCGCCCATGCTTG	3833
Db	541	LysAspAlaGlyLysLeuThrMetGlnProLeuValIleLeuAlaPheHisAlaVal	560

C:Genetics;
A:Gene: pol
C:Superfamily: pol polyprotein
F:1-120/Product: proteinase #status predicted <PTN>
F:3-102/Product: retropepsin #status predicted <RTP>
F:121-797/Product: RNA-directed DNA polymerase #status predicted <REV>
F:798-1189/Product: endonuclease #status predicted <EDB>
F:27/Active site: Asp (shared with dimeric partner) #status predicted

Alignment Scores:
Pred. No.: 1.63e-234 Length: 1189
Score: 3940.50 Matches: 758
Percent Similarity: 76.0% Conservative: 153
Best Local Similarity: 63.2% Mismatches: 261
Query Match: 26.9% Indels: 27
DB: 1 Gaps: 12

US-10-723-552-3 (1-8132) x GNMVW7 (1-1189)

Qy	2160	GGGAGACGGGTTCCGACCCCTCCCGAGCCGAGGTAACTTTGAAGTGGAGGGGCAA	2219
Db	1	GlyCysGlnGlySerGlyAlaProProGlnProArgLeuThrLeuSerValGlyGlyHis	20
Qy	2220	CCAGTTGAGTTCCTGGTTGATACCGGAGCGCAACATTCAGTGTCTACTACGCCATTAGGA	2279
Db	21	ProThrThrPheLeuValAspThrGlyAlaGlnHisSerValLeuThrLysAlaAsnGly	40
Qy	2280	AAACTAAAGATAAAAAATCCTGGGTGATGGTGCCACAGGCGCAACACAGTATCCATGG	2339
Db	41	ProLeuSerSerArgThrSerTrpValGlnGlyAlaThrGlyArgLysMetHisLysTrp	60
Qy	2340	ACTACCCGAAGAACAGTTGACTGGGAGTGGGACGGGTAAACCCACTGTTTCTGGTCATA	2399
Db	61	ThrAsnArgArgThrValAsnLeuGlyGlnGlyMetValThrHisSerPheLeuValVal	80
Qy	2400	CCTCAGTGCCAGCACCCCTCTAGGTAGACACTTATTGACCAAGATGGGAGCACAAATT	2459
Db	81	ProGluCysProTyrProLeuLeuGlyArgAspLeuLeuThrLysLeuGlyAlaGlnIle	100
Qy	2460	TCTTTT---GAACAAGGAAACCAAGAAGTGTCTGCAAAATACAAACCTTACCTGTGTG	2516
Db	101	HisPheSerGluAlaGlyAlaGlnValLeuAspArgAspGlyGlnProIleGlnIleLeu	120
Qy	2517	ACCTCCAAATTAGATGACGAATATCGACTACTCTCT---CCCTAGTTAAAG-----CCT	2567
Db	121	ThrValSerLeuGlnAspGluHisArgLeuPheAspIleProValThrThrSerLeuPro	140
Qy	2568	GATCAAAATATACAAATTCTGGTTGSMACAGTTTCCCAAGCTGGGAGAAACCGCAGGG	2627
Db	141	Asp-----ValTrpLeuGlnAspPheProGlnAlaTrpAlaGluThrGlyGly	156
Qy	2628	ATGGGTTTGGCAAGCAAGTTCCCGCCCAAGATTATCAACTGAAGGGCCAGTCCACACCA	2687
Db	157	LeuGlyArgAlaLysCysGlnAlaProIleIleLeuAspLeuLysProThrAlaValPro	176
Qy	2688	GTGTCAGTCAGACAGTACCCCTTGAGTAAAGAGCTCAAGGAATTCGGCCCGCATGTC	2747
Db	177	ValSerIleLysGlnTyrProMetSerLeuGluAlaHisMetGlyIleArgGlnHisIle	196
Qy	2748	CAAGATTAAATCCACAGGCGATCCTAGTTCCTGCTCCAAATCTCCCTGGTAATCTCCCTG	2807
Db	197	IleLysPheLeuGluLeuGlyValLeuArgProCysArgSerProTrpAsnThrProLeu	216
Qy	2808	CTACCGGTTAGAAAGCCCTGGGACTAATGACTATATCGACGATACAGGACTTCGAGAGATC	2867
Db	217	LeuProValLysLysProGlyThrGlnAspTyrArgProValGlnAspLeuArgGluIle	236
Qy	2868	AATAAACGGGTGAGGATATACCCCAACAGTCCCGCAACCCCTTAACCTCTTGTGTGCT	2927
Db	237	AsnLysArgThrValAspIleHisProThrValProAsnProTyrAsnLeuSerThr	256
Qy	2928	CTCCCAACCCCAACGGAGCTGGTATACAGTATTGGACTTTAAGGATGCTTCTCTGCTG	2987

Db	257	LeuLysProAspTyrSerTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeu	276
Qy	2988	AGATTACACCCCACTAGCCCAACACACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGA	3047
Db	277	ProLeuAlaProGlnSerGlnGluLeuPheAlaPheGluTrpLysAspProGluArgGly	296
Qy	3048	AGAACGGGCGAGCTCACCTCGACCCGACCTGCCCAAGGTTCAAGAACTCCCGACCATC	3107
Db	297	IleSerGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeu	316
Qy	3108	TTTGACGAAGCCCTTACACAGAGACCTGGCCCAACTTTCAGGATCCCAACACCTTCAGGTGACC	3167
Db	317	PheAspGluAlaLeuHisArgAspLeuThrAspPheArgThrGlnHisProGluValThr	336
Qy	3168	CTCCTCCAGTACGTGATGACCTGCTTCCTGGCGGAGCCACCAACAGAGACTCGCTAGAA	3227
Db	337	LeuLeuGlnTyrValAspAspLeuLeuAlaAlaProThrLysLysAlaCysThrGln	356
Qy	3228	GGCACGAAGGCATCTACTGCTGGAAATTGCTGACCTAGGCTACAGAGCTCTGCTAAGAAG	3287
Db	357	GlyThrArgHisLeuLeuGlnGluLeuGlyGlnLysGlyTyrArgAlaSerAlaLysLys	376
Qy	3288	GCCACAGATTTGCAGGAGAGAGGTAAACATCTTGGGGTACAGTTTTCGGGAGCGGCACCGA	3347
Db	377	AlaGlnIleCysGlnThrLysValThrTyrLeuGlyTyrIleLeuSerGluGlyLysArg	396
Qy	3348	TGCTGTACGAGGACCGAAGAAACTGTAGTCAGATACCGGCCCCCAACACAGCCAAA	3407
Db	397	TrpLeuThrProGlyArgIleGluThrValAlaArgIleProProArgAsnProArg	416
Qy	3408	CAATGAGACAGTTTTTGGGAGCAGCTGGATTTTGCAGACTGTGATCCCGGGTTTTCG	3467
Db	417	GluValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAla	436
Qy	3468	ACCTTAGCAGCCCATCTACTCCCGCTAAACCAAGAAAGGGAATTTCTCTGGGCTCCT	3527
Db	437	GluLeuAlaAlaProLeuTyrAlaLeuThrLysGluSerThrProPheThrTrpGlnThr	456
Qy	3528	GAGCACAGAGGATTTGATTCATAAAGCCCTGCTGAGCGCACCTGCTCTCGGCC	3587
Db	457	GluHisGlnLeuAlaPheGluAlaLeuLysLysAlaLeuLeuSerAlaProAlaLeuGly	476
Qy	3588	CTCCCTGACCTAACTAAACCTTTACCTTTTATGTGATGAGCTAAGGAGTAGGCCGG	3647
Db	477	LeuProAspThrSerLysProPheThrLeuPheLeuAspGluArgGlnGlyIleAlaLys	496
Qy	3648	GGAGTTTAAACCCAAACCTTAGGACCATGGAGAGAACCTGTCGCTACCTGTCAAAGAAG	3707
Db	497	GlyValLeuThrGlnLysLeuGlyProTrpLysArgProValAlaTyrLeuSerLysLys	516
Qy	3708	CTCGATCCTGTAGCCATGGTGGCCCATATGCTGGAAGGCTATCGAGCTGTGGCCATA	3767
Db	517	LeuAspProValAlaAlaGlyTyrProProCysLeuArgIleMetAlaAlaThrAlaMet	536
Qy	3768	CTGGTCAAGACCGCTGACAAATTGACTTTGGGACAGATATTAATGTAATAGCCCCCAT	3827
Db	537	LeuValLysAspSerAlaLysLeuThrLeuGlyGlnProLeuThrValIleThrProHis	556
Qy	3828	GCATTGGAGAACATCGTTTCGGCAGCCCCCAGACCGATGGATGACCAACCGCCGATGACC	3887
Db	557	ThrLeuGluAlaIleValArgGlnProProAspArgTrpIleThrAsnAlaArgLeuThr	576
Qy	3888	CATCATCAAGCCCTGCTTCTC---ACAGAGAGGCTACGTTTCGCTCCACCGCCGCTCTC	3944
Db	577	HisTyrGlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProProValThrLeu	596
Qy	3945	AACCTCGCCACTTCTGCTCGCTGAAGAGACTGATGAACCACTGACTCATGATGGCCATCA	4004
Db	597	AsnProAlaThrLeuLeuProValProGluAsnGlnProSerProHisAspCysArgGln	616
Qy	4005	CTATTGATTGAGGAGACTGGGGTCCCGCAAGGACCTTACAGACATACCGCTGACTGGAA	4064

Db 617 ValLeuAlaGluThrHisGlyThrArgGluAspLeuLysAspGlnGluLeuProAspAla 636
QY 4065 GTGCTAACCTGGTTCTACTGACGAGCAGCTATGTGTGGAGGTAAAGAGGATGGCTGGG 4124
Db 637 AspHisThrTrpTyrThrAspGlySerSerTyrLeuAspSerGlyThrArgAlaGly 656
QY 4125 GCGGGGTGGTGGAGCGGACCGCACATCTGGCCAGCAGCTCCGCGGAGGAAGTCA 4184
Db 657 AlaAlaValValAspGlyHisAsnThrIleTrpAlaGlnSerLeuProGlyThrSer 676
QY 4185 GCACAAAGGCTGAGCTCATGGCCCTCACGCAAGCTTTGGCGTGGCGGAGGGAATCC 4244
Db 677 AlaGlnLysAlaGluLeuIleAlaLeuThrLysAlaLeuGluLeuSerLysGlyLys 696
QY 4245 ATAAACATTATACGACAGCAGATATGCCCTTTCCGACTCGACACGTCACATGGGGCCATC 4304
Db 697 AlaAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAlaHisThrGlySerIle 716
QY 4305 TATAAACAAAGGGGTTCCTTACTCTACGAGGAGGGAATAAAGAAAGAGGAAT 4364
Db 717 TyrGluArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAlaGluIle 736
QY 4365 CTAAGCCTATTAGAAGCGGTACATTTACCAAAAAGGCTAGCTATTATACACTGCTCTGGA 4424
Db 737 IleAlaLeuLeuLysAlaLeuPheLeuProGlnGluValAlaIleIleHisCysProGly 756
QY 4425 CATCAGAAAGCTAAAGATCTCATATCCAGAGAGAAACAGATGGCTGACCGGGTTGCCAAG 4484
Db 757 HisGlnLysGlyGlnAspProValAlaValAlaGlyAsnArgGlnAlaAspArgValAlaArg 776
QY 4485 CAGCAGCCACCGGTGTAACTCTCTCCCTATATAGAAATGCCAAGCCCCAGAACCC 4544
Db 777 GlnAlaAla-----MetalGluValLeuThrLeuAlaThrGluProAspAsn 792
QY 4545 -----AGACGACAGTACACCTCAGACACTGGCAAGAGATAAAGAG 4586
Db 793 ThrSerHisIleThrIleGluHisThrTyrThrSerGluAspGlnGluAlaAlaArgAla 812
QY 4587 ATAGACAGTTCTTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGGAATC 4646
Db 813 IleGlyAla--ThrGluAsnLysAspThrArgAsnTrpGluLysGluGlyLysIleVal 831
QY 4647 CTCGCCCAAGAGAGGTTAGATATGTCACAGATATACGTCTAACCCACTTAGGA 4706
Db 832 LeuProGlnLysGluAlaLeuAlaMetIleGlnMetHisAlaTrpThrHisLeuGly 851
QY 4707 ACTAAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATCTTCTGAGCTACCAGGA 4766
Db 852 AsnArgLysLeuLysLeuIleGluLysThrAspPheLeuIleProArgAlaSerThr 871
QY 4767 GTGGCTGACTCGGTGCTCAACATTTGTGCTCCCTGCGCAGCTGTTATGTAATCTCTTCC 4826
Db 872 LeuIleGluGlnValThrSerAlaCysLysValCysGlnGlnValAsnAlaGlyAlaThr 891
QY 4827 AGAATGCTCCAGGAGAGACTAAGGGGAAGCCACCAGCGCTCACTGGGAAGTGGAC 4886
Db 892 ArgValProAlaGlyLysArgThrArgGlyAsnArgProGlyValTyrTrpGluIleAsp 911
QY 4887 TTCACGTAGGTAAAGCCGGCTAATACGGGAACAAATACCTATTGCTTTTGTAGACACC 4946
Db 912 PheThrGluValLysProHisTyrAlaGlyTyrLysTyrLeuLeuValPheValAspThr 931
QY 4947 TTTTCAGGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 5006
Db 932 PheSerGlyTrpValGluAlaPheProThrArgGlnGluThrAlaHisIleValAlaLys 951
QY 5007 AAATATCTGGAAGAAATTTTCCAGATTGGGAATACCTAAGGTAAATAGGGTCAGACAAT 5066
Db 952 LysIleLeuGluGluIlePheProArgPheGlyLeuProLysValIleGlySerAspAsn 971
QY 5067 GGTCCAGCTTTTCTCCCGAGGTAACTCAGGAGCTGGCCAGATATATGGGATGATTGG 5126
Db 972 GlyProAlaPheValSerGlnValSerGlnGlyLeuAlaArgIleLeuGlyIleAsnTrp 991

QY 5127 AAATGTCATTGTGCATACAGACCCCAAGCTCAGGACGAGTAGAGAGTGAATAGAACCC 5186
Db 992 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 1011
QY 5187 ATTAAGAGACCTTACTTAATTAATGACCGCGAGACTGCGCTTAATGATTGGATAGCTCTC 5246
Db 1012 IleLysGluThrLeuThrLysLeuGluThrLeuGluThrGlyLeuLysAspTrpArgLeu 1031
QY 5247 CTCGCCCTTTGCTGCTTTTAGGGTTAGGAACACCCCTGGACAGCTTGGGCTGACCCCTAT 5306
Db 1032 LeuSerLeuAlaLeuAlaArgAlaArgAsnThrProAsnArgPheGlyLeuThrProTyr 1051
QY 5307 GAATTAATCTTACCGGGGACCCCCCATTCGTAGAAAATTCCTCTGTACATAGTGTCTCAC 5366
Db 1052 GluIleLeuTyrGlyGlyProProLeuSerThrLeuLeuAsnSerPheSerProSer 1071
QY 5367 GTCTGCTTCTCCAGCCTTTGCTCTAGGCTCAAGGCACTTGTAGTGGGTGAGACAACGA 5426
Db 1072 AsnSerLysThrAsp---LeuGlnAlaArgLeuLysGlyLeuGlnAlaValGlnAlaGln 1090
QY 5427 GCGTGGAGGCACTCCGGGAGGCTACTCAGGAGGAGGACTTGCAGATCCCACTCGT 5486
Db 1091 IleTrpAlaProLeuAlaGluLeuTyrArgProGlyHisSer---GlnThrSerHisPro 1109
QY 5487 TTCCAAAGTGGGATTCAGTCTACGTTAGACGCCACCGTCGAGMAACCTCGAGACTCGG 5546
Db 1110 PheGlnValGlyAspSerValTyrValArgArgHisArgSerGlnGlyLeuGluProArg 1129
QY 5547 TGAAGGGCCCTTATCTCGTACTTTTGACACACCAACCGCTGTGAAGTCTGAAGGAATC 5606
Db 1130 TrpLysGlyProTyrIleValLeuLeuThrThrProThrAlaIleLysValAspGlyIle 1149
QY 5607 TCACCTGGATCCATCCACCTCCACGTTAAACCGCGCCCA-----CCTCCCGAT 5654
Db 1150 AlaThrTrpIleHisAlaSerHisAlaLysAlaAlaProGlyThrProGlyProThrSer 1169
QY 5655 TCGGGG---TGGAAAGCCGAAAGACTGAAATATCCCTTAAAGCTTGCCTCCATCCG 5708
Db 1170 SerGlyThrTrpArgLeuArgArgSerGluAspProLeuLysIleArgLeuSerArg 1188

RESULT 9
GNMVC
pol polyprotein - feline endogenous virus BCE1 (fragment)
N:Contains: endonuclease (EC 3.1.-.-); RNA-directed DNA polymerase (EC 2.7.7.49)
C:Species: feline endogenous virus BCE1
A:Note: host Felis silvestris catus (domestic cat)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S12813; S12814
R:Moehring, R.
submitted to the EMBL Data Library, February 1990
A:Reference number: S12812
A:Accession: S12813
A:Molecule type: DNA
A:Residues: 1-1046 <MO>
A:Cross-references: UNIPROT:P31792; UNIPARC:UPI0000131ED7; EMBL:X51929
C:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: hydrolase; nucleotidyltransferase; polyprotein
F;1-647/Product: RNA-directed DNA polymerase (fragment) #status predicted <RTP>
F;648-1046/Product: endonuclease #status predicted <EDC>

Alignment Scores:
Pred. No.: 118e-209 Length: 1046
Score: 3536.50 Matches: 676
Percent Similarity: 76.4% Conservative: 129
Best Local Similarity: 64.1% Mismatches: 228
Query Match: 24.2% Indels: 21
DB: 1 Gaps: 8

US-10-723-552-3 (1-8132) x GNMVCE (1-1046)


```
Db 718 LysThrAspPheLeuLeuProLysAlaGlyThrLeuLeuGluGlnValThrSerAlaCys 737
Qy 4794 GTGCCCTCCAGCTGGTTAAATCCTTCCAGAAATGCTCCAGGAGAGACTAAGG 4853
Db 738 LysValCysGlnGlnValAsnAlaGlyAlaThrArgValProGluGlyLysArgThrArg 757
Qy 4854 GGAAGCCACCCAGCGCTCACTGGAAGTGAAGTCACTGAGTAAAGCCGGCTAAATAC 4913
Db 758 GlyAsnArgProGlyValTyrTrpGluLeuAspPheThrGluValLysProHisTyrAla 777
Qy 4914 GGAACCAATACCTATTGGTTTGTAGACACCTTTTCAGGATGGGTAGAGCTTATCCT 4973
Db 778 GlyTyrLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyrPro 797
Qy 4974 ACTAAGAAAGAGACTTCAACCGTGGCTGCTTAAAAAATCTGGAAGAAATTTTCCAAGA 5033
Db 798 ThrArgGlnGluThrAlaHisMetValAlaLysLysIleLeuGluGluIlePheProArg 817
Qy 5034 TTTGGAATACCTAAGTAATAGGTCAGACAATGGTCCAGCTTTTGTGGCCAGGTAAAT 5093
Db 818 PheGlyLeuProLysValIleGlySerAspAsnGlyProAlaPheValSerGlnValSer 837
Qy 5094 CAGGACTGGCCAAAGATATTGGGATTGATTGGAAATCTGCAATTCGTGCATACAGACCCCAA 5153
Db 838 GlnGlyLeuAlaArgThrLeuGlyLysIleAsnTrpLysLeuHisCysAlaTyrArgProGln 857
Qy 5154 AGTCAGGACAGGTAGAGAGATGAATAGAACCAATTAAGAGACCCCTTACTAAATGACC 5213
Db 858 SerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThr 877
Qy 5214 GCGGAGACTGGCTTAATGATTGATAGTACTCTCTGCGCTTGTGCTTTTGGTTTGGT 5273
Db 878 LeuGluThrGlyLeuLysAspTrpArgLeuLeuSerLeuAlaLeuLeuArgAlaArg 897
Qy 5274 AACACCCCTGGACAGTTTGGCTGACCCCTATGAATTAATCTCTACGGGGGACCCCCCA 5333
Db 898 AsnThrProAsnArgPheGlyLeuThrProTyrGluLeuLeuTyrGlyLysProPro 917
Qy 5334 TTGGTAAATTTGCTCTCATAGTGTGACGTGCTGCTGCTTTTCCAGCCT- 5384
Db 918 LeuSerThrLeuLeuAsnSerPheSerPro- 5384
Qy 5385 TTGTTCTTAGGCTCAAGGCACTTGAGTGGGTGAGACAGAGCGTGGAGGCACTCCGG 5444
Db 934 LeuGlnAlaArgLeuLysGlyLeuGlnAlaValGlnAlaGlnIleTrpThrProLeuAla 953
Qy 5445 GAGSCCTACTCAGGAGGAGGAGCTTCAGATCCACATCGTTTCCAGTGGGAGTCA 5504
Db 954 GluLeuTyrArgProGlyHisPro- 5504
Qy 5505 GTCTACGTTAGACGCCACCGTCAGGAAACCTCGAGACTCGGTGGAAGGCGCTTATCTC 5564
Db 973 ValTyrValArgTrpHisArgSerGlnGlyLeuGluProArgTrpLysGlyProTyrIle 992
Qy 5565 GTACTTTTGACCAACCAAGCGGTGGAAGTGAAGTGAAGTTCCTGATCCATGCA 5624
Db 993 ValLeuLeuThrThrProThrAlaIleLysValAspGlyIleAlaAlaTrpIleHisAla 1012
Qy 5625 TCCACGCTTAA- 5625
Db 1013 SerHisAlaLysAlaProLysThrProGlyProGluThrProLysThrTrpLysLeu 1032
Qy 5670 GAAAGACTGAAATCCCTTAAAGCTTCGCTCCATCGCGTG 5711
Db 1033 HisArgSerGluAsnProLeuLysIleArgLeuSerArgVal 1046
```

RESULT 10

GNVWK

pol polyprotein - AKR murine leukemia virus (fragment)

C;Species: AKR murine leukemia virus

C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004

C;Accession: A03958

R;Herr, W.; Corbin, V.; Gilbert, W.

Nucleic Acids Res. 10, 6931-6944, 1982
A;Title: Nucleotide sequence of the 3' half of AKV.
A;Reference number: A93448; MUID:83090450; PMID:6294621
A;Accession: A03958
A;Molecule type: DNA
A;Residues: 1-843 <HR>
A;Cross-references: UNIPROT:P03357; UNIPARC:UPI00000131P01
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Alignment Scores:
Pred. No.: 1.86e-154 Length: 843
Score: 2639.50 Matches: 504
Percent Similarity: 75.5% Conservative: 133
Best Local Similarity: 59.7% Mismatches: 194
Query Match: 18.0% Indels: 13
DB: 1 Gaps: 8

US-10-723-552-3 (1-8132) x GNVWK (1-843)

```
Qy 3216 GACTGCTTAGAGGCGACGAGGCACTACTCTGGAATTGTCTGACCTAGGCTACAGACC 3275
Db 3 AspCysGlnGlnGlyThrArgAlaLeuLeuLeuThrLeuGlyAsnLeuGlyTyrArgAla 22
Qy 3276 TCTGCTAAGAGGCGCCAGATTTCAGGAGAGAGCTAACTACTTGGGGTACAGTTTCGG 3335
Db 23 SerAlaLysLysAlaGlnLeuCysGlnLysGlnValLysTyrLeuGlyTyrLeuLys 42
Qy 3336 GACGGGCGAGTGGCTGACGGAGGCGACGGAAGAAAATCTAGTCCAGATACCGGCCCA 3395
Db 43 GluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGlyGlnProThrPro 62
Qy 3396 ACCACAGCCAAAATAATGAGAGAGATTTCGGGACAGCTGGATTTTCAGACTGTGATC 3455
Db 63 LysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIle 82
Qy 3456 CCGGGGTTTCGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAAGGGGAATTC 3515
Db 83 ProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPhe 102
Qy 3516 TCCTGGGCTCTGACGACCCAGAGGCATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCA 3575
Db 103 AsnTrpGlyProArgGlnGlnLysAlaTyrGlnLysGlnAlaLeuLeuThrAla 122
Qy 3576 CCGTCTGCGCCCTCCCTGACGTAACTAAACCTTTTACCTTTATGTGGATGAGCGTAAG 3635
Db 123 ProAlaLeuGlyLeuProAspLeuThrLysPheGluLeuPheValAspGluLysGln 142
Qy 3636 GGAAGTACCGCGGGAGTTTAAACCAACCTTAGGACCATGGAGAGACCTGTGCGCTAC 3695
Db 143 GlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArgArgProValAlaTyr 162
Qy 3696 CTCTCAAGAGCTCGATCCTGTAGCCAGGTGGTGGCCCAATATGCTTGAAGGCTATCGCA 3755
Db 163 LeuSerLysLysLeuAspProValAlaAlaGlyTrpProProCysLeuArgMetValAla 182
Qy 3756 GCTGTGCCATATCTGTCAGGACGCTGACAAATTTGACTTTGGGACAGATATACTGTA 3815
Db 183 AlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIle 202
Qy 3816 ATAGCCCCCATCATTTGGAGAACATCGTTTCGGCAGCCCCCAGACCGATGGATGACCAAC 3875
Db 203 LeuAlaProHisAlaValGluAlaLeuValLysGlnProAspArgTrpLeuSerAsn 222
Qy 3876 GCCCGCATGACCCATATCAAAGCCTGCTTCTC---ACAGAGGGGTCACTGTTGCTCCA 3932
Db 223 AlaArgMetThrHisTyrGlnAlaMetLeuLeuAspThrAspArgValGlnPheGlyPro 242
Qy 3933 CAGCGCGCTCTAACCCCTGCGCTCTCTGCTGGAAGAGACTGATGAACAGTGTGCTCAT 3992
Db 243 ValValAlaLeuAsnProAlaThrLeuLeuPro---LeuProGluGluGlyAlaProHis 261
```


DB:	2	Gaps:	9
US-10-723-552-3	(1-8132) x A46311	(1-559)	
QY	1476	TTTCCCTTAACCTCCGCCCGTTGGACTACAAACGGCTGAAGGTAGGAGAGCTTGAAA	1535
Db	2	PheProLeuGluArgProAspGlyThrGlnAlaGlyArgSerProThrAsnHisLeuVal	21
QY	1536	ATCTATATCGCCAGGCTCTGGTGGCGGGTCTCCGGGCGCCTCAAGACGCGCCCACTAAATTG	1595
Db	22	HisTyrArgGlnLeuLeuAlaGlyLeuGlnAsnAlaGlyArgSerProThrAsnLeu	41
QY	1596	GCTAAGGTACAGAAGTATGTCAGGAGCCGATGAACCCCTCTGTTTCTTGTAGAGG	1655
Db	42	AlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSerAlaPheLeuGluArg	61
QY	1656	CTCTTGAAGCCTTCAGSCGTACACCCCTTTTGATCCACCTCAGAGGCCCAAAAGCC	1715
Db	62	LeuLysGluAlaTyrArgTyrThrProTyrAspProGluAspProGlyGlnGluThr	81
QY	1716	TCAGTGCCTTTGGCCTTTATAGACAGCTCAGCCCTTGCATATTAGAAAGAAAGCTTCAGAGA	1775
Db	82	AsnValSerMetSerPheIleTyrGlnSerAlaProAspIleGlyArgLysLeuGlyArg	101
QY	1776	CTGAAGGGTTACAGAGGCTGATTAAGTGTACTAGTGAAGGAGGACAGAGAAGTATAT	1835
Db	102	LeuGluAspLeuLysSerLysThrLeuGlyAspLeuValArgGluAlaGluLysIlePhe	121
QY	1836	TACAAAAGGGAGACAGAAGAAGGAAGGACAAAGAAAGAGAGAGAAAGAGAGGAAGG	1895
Db	122	AsnLysArgGluThrProGluGluArgGluArgIleArgArgGluThrGluGluLys	141
QY	1896	GAGGAAGA-----GCTAATAAACCGCAAGG	1922
Db	142	GluGluArgArgThrValAspGluGlnLysGluArgAspArgArgArgHis	161
QY	1923	AAGAATTGACTAAGATCTTGGCTGCAGTGGTTGAAGGGAAAGCAATACGGAAGAGAG	1982
Db	162	ArgGluMetSerLysLeuLeuAlaThrValValIleGly-----GlnGlu	176
QY	1983	AGAGATTTAGGAAAATTAGGTCAAGGCCCTAGACAGTCAGGAGCACTGGGCAATAGGACC	2042
Db	177	GlnAsp-----ArgGlnGluGlyGlu---ArgLysArgPro	187
QY	2043	CCACTCGACAAGGACCAATGTGCATATTGTAAAGAAAGAGACACTGGGCAAGCACTGC	2102
Db	188	GlnLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaLysAspCys	207
QY	2103	CCCAAGAAG-----GGAAACAAAGGACCAAGG-----ATCCTAGCTCTAGAA	2144
Db	208	ProLysLysProArgGlyProArgLysProArgProGlnThrSerLeuLeuThrLeuGly	227
QY	2145	GAAGATAAAGATTAGGGAGACGGGTTTCGGACCCCTCCCGAGCCCGGTAACTTTG	2204
Db	228	Asp-----GlyGlyGlnGlyGlnAspProProGluProArgIleThrLeu	243
QY	2205	AAGGTGAGGGGCAACCAAGTTGATGCTCTGTTGATACCGGAGCGAAACATTCCAGTGCTA	2264
Db	244	LysValGlyGlyGlnProValThrPheLeuValAspThrGlyAlaGlnHisSerValLeu	263
QY	2265	CTACAGCCATTAGGAAAACATAAAGATAAAAAATCTGGGTGATGGGTGCCAGGGCCAA	2324
Db	264	ThrGlnAsnProGlyProLeuSerAspLysSerAlaTrpValGlnGlyAlaThrGlyGly	283
QY	2325	CAACAGTATCCATGGACTACCCGAAGAACAGTTGACTTGGGATGGGACGGGTAAACCCAC	2384
Db	284	LysArgTyrArgTrpThrThrAspArgLysValHisLeuAlaThrGlyLysValThrHis	303
QY	2385	TCGTTTCTGTCATCTAGTCCAGCACCCCTCTTAGGTAGAGACTTATTGACCAAG	2444
Db	304	SerPheLeuHisValProAspCysProTyrProLeuLeuGlyArgAspLeuLeuThrLys	323
QY	2445	ATGGGAGCACAAAATTTCTTTTGAA-----CAAGGGAAACCAAGAGTGTCTGCAATAACAAA	2501

Percent Similarity: 76.6%		Conservative: 49
Best Local Similarity: 66.9%		Mismatches: 92
Query Match: 11.8%		Indels: 25
DB: 2		Gaps: 4
US-10-723-552-3 (1-8132) x S04842 (1-476)		
QY	3441	TGCAGCTGTCGATCCCGGGTTCGACCTTAGCAGCCCTCCTACCCGCTAACCAAA 3500
Db	1	CysargLeuTrpIleProGlyPheAlaSerLeuAlaProLeuTyrProLeuThrLys 20
QY	3501	GAATAAGGGGAATTCCTCCGTGGCTCCTGAGCACCAGAGGCAATTTGATGCTATCAAAAAG 3560
Db	21	GluserIleProPheIleTrpThrGluGluHisGlnLysAlaPheAspArgIleLysGlu 40
QY	3561	GCCTGCTGAGCGCAGCTGCTCTGGCTCCCTGAGCGTAATAACCTTTACCTTTAT 3620
Db	41	AlaLeuLeuSerAlaProAlaLeuAlaLeuProAspLeuAlaLysProPheThrLeuTyr 60
QY	3621	GTGGATGAGCTAAGGAGTAGCCCGGGAGCTTTTAACCCAAACCTAGGACCATGGAGA 3680
Db	61	ValAspGluArgAlaGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArg 80
QY	3681	AGACTGTGCTGCTACCTCTCAAGAAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGC 3740
Db	81	ArgProValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProThrCys 100
QY	3741	CTGAAGCTATCGCAGCTGTGGCCATCTGCTCAAGGACGCTGACAAATTCATTTGGGA 3800
Db	101	LeuLysAlaValAlaValAlaLeuLeuLeuLysLysPheAlaAspLysLeuThrLeuGly 120
QY	3801	CAGATATACTGTAATAGCCCCCATCATCTGATGGAGAACATCGTTCGGCAGCCCCAGAC 3860
Db	121	GlnAsnValThrValIleAlaSerHisSerLeuGluSerIleValArgGlnProProAsp 140
QY	3861	CGATGGATGACCAACGCCCGCATGACCCACTATCAAGCCTGCTTCTCACAGAGGGTCTC 3920
Db	141	ArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuAsnGluArgVal 160
QY	3921	ACGTTGCTGCCACGACGGCTCTCAACCTGCCACTCTTCCTGCTGAAGACTGATGAA 3980
Db	161	SerpPheAlaProAlaValLeuAsnProAlaThrLeuLeuProValGluSerGlnAla 180
QY	3981	CCAGTGACTCATGATGCTCACTCACTATGATTGAGGAGACTGGGGTCCGCAAGACCTT 4040
Db	181	ThrProValHisArgCysSerGluIleLeuAlaGluGluThrGlyThrArgArgAspLeu 200
QY	4041	ACAGACATACCGCTGACTGGAGAACTGTAACCTGTTCACTGACGGAAAGCAGCTATGTG 4100
Db	201	LysAspGlnProLeuProGly---ValProAlaTrpTyrThrAspGlySerSerPheIle 219
QY	4101	GTGAAGGTAAAGAGATCGTGGGCGCGGTGTGGACGGGACCCGACCATCTGGGCC 4160
Db	220	AlaGluGlyLysArgArgAlaGlyAlaAlaIleValAspGlyLysArgThrValTrpAla 239
QY	4161	AGCAGCTTCGGGAGGAACTTCAGCAAAAGCTGAGCTCATGGCCCTCACCCAGCT 4220
Db	240	SerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuValAlaLeuThrGlnAla 259
QY	4221	TTGGCGCTGGCGGAGGAAATCCATAAATTTATACGGACAGCAGGTATGCTCTTTCG 4280
Db	260	LeuArgLeuAlaGluGlyArgAspIleAsnIleTyrThrAspSerArgTyrAlaPheAla 279
QY	4281	ACTGCACAGTACATGGGGCCATCTATAAACAAGGGGTGCTTTACCTCAGCAGGGAGG 4340
Db	280	ThrAlaHisIleHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyLys 299
QY	4341	GAATAAAGAACAAAGAGCAATCTAAGCTTATTAGACCGGTACATTTACCCAAAGG 4400
Db	300	AspIleLysAsnLysGluGluLeuAlaLeuLeuGluAlaIleHisLeuProLysArg 319
QY	4401	CTAGCTATTATACACTGTCCTGGACATCAGAAAAGCTAAAGATCTCATATCCAGAGGAAC 4460

Db	320	ValAlaIleIleHisCysProGlyHisGlnLysGlyAsnAspProValAlaThrGlyAsn 339
QY	4461	CAGATGCTGCTACCGGGTTCGACGAGCCAGCCAGGGGTGTTAACTCTTCCTCTATAATA 4520
Db	340	ArgArgAlaAspGluAlaAlaLysGlnAlaLeuSerThrArgValLeuAlaGluThr 359
QY	4521	GAATGCCCCAAAGCCCAAGAACCCAGACGACATCACCTTAGAAGACTGCGCAAGAGATA 4580
Db	360	ThrLysPro-----GlnGluLeu 365
QY	4581	AAAAAGATAGACCACTTCTGTGAGACTCCGGAAGGACCTGCTATATCCTCAGATGGGAAG 4640
Db	366	IleLysProAlaGlnValLysThrArgPro-----Gly 376
QY	4641	GAATCTGCTGCCCCAACAAAGAGGGTTAGATATGTCACACAGATACATCGTCAACCCAC 4700
Db	377	GluLeuThrProAspArg---GlyLysGluPheIleGlnArgLeuHisGlnLeuThrHis 395
QY	4701	CTAGGAATAAACACCTCTGCAGCAGTTCGTTCAGAACATCCCCCTTATCATGTTCTCAGGCTA 4760
Db	396	LeuGlyProGluLysLeuLeuGlnLeuValAsnArgThrSerLeuLeuLeuProAsnLeu 415
QY	4761	CCAGAGTGGCTGACTCGGTGTCAAACATTTGTGTCGCCAGCTGGTTAATGCTAAT 4820
Db	416	GlnSerAlaValArgGluValThrSerGlnCysGlnAlaCysAlaMetThrAsnAlaVal 435
QY	4821	CCTTCCAGATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCCTCTGGGAA 4880
Db	436	ThrThrTyrArgGluThrGlyLysArgGlnArgGlyAspArgProGlyValTyrTrpGlu 455
QY	4881	GTGACATCTCACTGAGGTAAAGCCGCTAAATACCGAAACAAATACCTATTGGTTTTGTA 4940
Db	456	ValAspPheThrGluValLysProGlyArgTyrGlyAsnArgTyrLeuLeuValPheIle 475
QY	4941	GAC 4943
Db	476	Asp 476
RESULT 13		
A42743		
pol polyprotein - radiation murine leukemia virus (strain Kaplan) (fragment)		
N;Contains: endonuclease (EC 3.1.1.-); proteinase (EC 3.4.21.-); RNA-directed DNA polym		
C;Species: radiation murine leukemia virus		
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004		
C;Accession: A42743		
R;Poliquin, L.; Bergeron, D.; Fortier, J.L.; Paquette, Y.; Bergeron, R.; Rassart, E.		
A;Title: Determinants of thymotropism in Kaplan radiation leukemia virus and nucleotide		
A;Reference number: A42743; MUID:92333703; PMID:1629969		
A;Accession: A42743		
A;Molecule type: DNA		
A;Residues: 1-581 <POL>		
A;Cross-references: UNIPROT:P31795; UNIPARC:UPI0000131P09; GB:M93052; NID:9332065; PIDN:		
C;Comment: This protein is likely to be expressed as a fused gag-pol polyprotein.		
C;Comment: The precise boundary between the mature proteins has not been determined.		
C;Genetics:		
A;Gene: pol		
C;Superfamily: pol polyprotein		
C;Keywords: endonuclease; hydrolase; nucleotidyltransferase; polyprotein; reverse trans		
Alignment Scores:		
Pred. No.:	1.73e-95	Length: 581
Score:	1681.00	Matches: 329
Percent Similarity:	69.9%	Conservative: 84
Best Local Similarity:	55.7%	Mismatches: 148
Query Match:	11.5%	Indels: 30
DB:	2	Gaps: 6
US-10-723-552-3 (1-8132) x A42743 (1-581)		
QY	4002	CAACTATTGATGTAGAGACTGGGGTCCGGAAGGACCTTACAGACATACCCCTGACTCGA 4061
Db	2	GluIleLeuAlaGluThrHisGlyThrArgProAspLeuThrAspGlnProIleProAsp 21

Query Match:	11.5%	Indels:	67
DB:	1	Gaps:	10
US-10-723-552-3 (1-8132) x FOLJGL (1-520)			
Qy	585	ATGGACAGAGCGGTGACGACCCCTCTTAGTTTGACTCTCGACCAATTGAGCTGAAGTTAAA	644
Db	1	MetGlyGlnAspAenSerThrProIleSerLeuThrLeuAenHisTrpArgAspValArg	20
Qy	645	TCAGGGCTCATATTTGTCAGTTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC	704
Db	21	ThrArgAlaHisAenLeuSerValGluIleLysGlyLysTrpGlnThrPheCysSer	40
Qy	705	TCTGAATGGCGACATTCGATGTTGGATGGCCATCAGAGGGACCTTAAATTCGAGATT	764
Db	41	SerGlnTrpProThrPheGlyValGlyTrpProGluGlyThrPheAenLeuSerVal	60
Qy	765	ATCTGGCTGTTAAAGCAGATTATTTTCAGACTGGACCCGGCTCTCATCCGATCAGGAG	824
Db	61	IlePheAlaValLysLysIleValPheGlnGluAen---GlyGlyHisProAspGlnVal	79
Qy	825	CCCTATATCTTCAGTGGCAGAGATTGGCAGAGATCTCCGCCATGGTT-----	875
Db	80	ProTyrIleValValTrpGlnAspLeuAlaGlnAenProProTrpValProAlaSer	99
Qy	876	-----AAACCATGGCTGATTAAGCCCAAGAAAGCCA	905
Db	100	AlaLysValAlaValSerAspThrArgArgProValAlaGlyArgProSerAlaPro	119
Qy	906	GGTCCCCGAATTCCTGGCTCTTGGAGAGAAAAACAACACTCGGCTGAAAAAGTCAAGCCC	965
Db	120	-----Pro	120
Qy	966	TCTCTCATATCTACCCGAG-----ATTGAGGAGCCACCGGCTTGG	1007
Db	121	ArgProProIleTyrProAlaThrAspAspLeuLeuLeuSerGluProThrProPro	140
Qy	1008	CCGGAACCCCAATCTGTTCCTCCACCCCTTATCTGGCAGACGGTGGCGGAGGGACCC	1067
Db	141	ProTyrProAlaAlaLeuProProPro-----LeuAlaProGlnAlaIleGlyProPro	158
Qy	1068	TTTCCCTCT---CCTGGAGCTCCGCGGTGGAGGACCTGTCTGAGGAGACTCGGAGCGG	1124
Db	159	SerGlyGlnMetProAspSerSerAspProGluGlyProAlaAlaGlyThrArgSerArg	178
Qy	1125	AGGGGCGCCACCCG-----GAGGGACAGACGAGATCGCGACATTACCGCTGGC	1175
Db	179	ArgAlaArgSerProAlaAspAsnSerGlyProAspSerThrValIleLeuProLeuArg	198
Qy	1176	ACGTACGGCTCCACACCGGGGGCCAATTGCGCCCTCCAGATTCGCCCTTTCT	1235
Db	199	AlaIleGlyProProAlaGluProAsnGlyLeuValProLeuGlnTyrTrpProPheSer	218
Qy	1236	TCTCAGATCTCTATAATTGAAAACTAACCATCCCTTTCTCGGAGGATCCCAACGC	1295
Db	219	SerAlaAspLeuTyrAenTrpLysSerAenHisProSerPheSerGluAenProAlaGly	238
Qy	1296	CTCAGGGGTGGTGGAGTCCCTTATGTTCTCACCAGCCTACTTGGATGATGTCAA	1355
Db	239	LeuThrGlyLeuLeuGluSerLeuMetPheSerHisGlnProThrTrpAspCysGln	258
Qy	1356	CAGCTGTGCGACACTCTTCACACCGGAGCGAGAGAATCTATTAGAGGCTAGA	1415
Db	259	GlnLeuLeuGlnIleLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArg	278
Qy	1416	AAAAATGTTCTGGGGCGCAGCGGCGACCCACGCGGTTGCAAAATGAGATTGACATGGGA	1475
Db	279	LysAenValLeuGlyAspAenGlyAlaProThrGlnLeuGluAenLeuIleAenGluAla	298
Qy	1476	TTTCCCTTAATCTCGCCCGGTGGGACTTACAAACCGCTGAAGTAGGAGAGCTTGAAA	1535
Db	299	PheProLeuAenArgProHisTrpAspTyrAenThrAlaAlaGlyArgGluArgLeuLeu	318

Qy	1536	ATCTATCCGACGCTCTGTGTGGGGGTCTCCGGGGCGCCTCAAGACGCCCTCAATTTG	1595
Db	319	ValTyrArgArgThrLeuValAlaGlyLeuLysGlyAlaAlaArgArgProThrAenLeu	338
Qy	1596	GCTAAGGTAAAGAGATGTCAGGGACCGAATGAACCCCTCTGTTTCTTTGAGAGG	1655
Db	339	AlaLysValArgGluValLeuGlnGlyProAlaGluProSerValPheLeuGluArg	358
Qy	1656	CTCTTGAAGCCCTCAGGCGGTACACCCCTTTGATCCCACTCAGAGGCCCAAAAGCC	1715
Db	359	LeuMetGluAlaTyrArgArgTyrThrProPheAspProSerSerGluGlyGlnAla	378
Qy	1716	TCAGTGGCTTTGGCTTTATAGACACTCAGCTTCGATATTAGAAAGAGCTTCAGAGA	1775
Db	379	AlaValAlaMetAlaPheIleGlyGlnSerAlaProAspIleLysLysLeuGlnArg	398
Qy	1776	CTGGAAGGTTTACAGGAGGCTGAGTTACGTGATCTAGTGAAGGAGGAGAAAGTATAT	1835
Db	399	LeuGluGlyLeuGlnAspTyrSerLeuGlnAspLeuValLysGluAlaGluLysValTyr	418
Qy	1836	TACAAAGGAGACAGAGAAAGGAAAGGAAACAAAGAAAGAGAGAGAGAGAAAGG	1895
Db	419	HisLysArgGluThrGluGluGluArgGlnGluArgGlnLysGluAlaGluGluLys	438
Qy	1896	GAGAAAGACGTAATAACCGCAAGAGAAGAAATTTGACTAAGATCTTGGCTGCACTGTT	1955
Db	439	GluArgArgArgAspArgProLysLysLysAenLeuThrLysIleLeuAlaValVal	458
Qy	1956	GAAGGAAAGCAATACGGAAGAGAGAGATTTTAGGAAATTTAGGTACGGCCCTAGA	2015
Db	459	SerArgGluGlySerThr-----GlyArg	466
Qy	2016	CAGTCAGGGAACCTGGCGAAT-----AGGACCCCACTC	2048
Db	467	GlnThrGlyAenLeuSerAenGlnAlaLysLysThrProArgAspGlyArgProLeu	486
Qy	2049	GACAAGACCAATGTGATATTGTAAAGAAAGAGACACTGGGCAAGAACTGCCCAAG	2108
Db	487	AspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaArgGluCysProArg	506
Qy	2109	AAGGAAAC---AAAGACCAAGGATCTAGCTTAGAA	2144
Db	507	LysLysHisValArgGluAlaLysValLeuAlaLeuAsp	519
RESULT 15			
T09394			
gag-pro-pol polyprotein - walleye dermal sarcoma virus			
C;Species: walleye dermal sarcoma virus			
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004			
C;Accession: T09394; T09393			
R;Petropoulos, C.J.			
submitted to the EMBL Data Library, November 1997			
A;Description: Appendix 2: Retroviral taxonomy, protein structure, sequences, and genet			
A;Reference number: 216660			
A;Accession: T09394			
A;Status: translated from GB/EMBL/DBJ			
A;Molecule type: genomic RNA			
A;Residues: 1-1751 <PE>			
A;Cross-references: UNIPROT:O92815; UNIPARC:UPI000010BDA0; EMBL:AF033822; NID:g2801519;			
A;Accession: T09393			
A;Status: translated from GB/EMBL/DBJ			
A;Molecule type: genomic RNA			
A;Residues: 1-582 <PEW>			
A;Cross-references: UNIPARC:UPI000010499D; EMBL:AF033822; NID:g2801519; PID:g2801521			
C;Genetics:			
A;Gene: gag-pro-pol			
A;Introns: 582/3			
Alignment Scores:			
Pred. No.: 5.2e-93 Length: 1751			
Score: 1643.00 Matches: 489			
Percent Similarity: 43.5% Conservative: 251			
Best Local Similarity: 28.7% Mismatches: 681			

Query Match:	11.2%	Indels:	280
DB:	2	Gaps:	42
US-1023-552-3 (1-8132) x T09394 (1-1751)			
QY	939	AAACACTCGGTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGAGCCA	998
DB	177	LysArgGluLysLysLysLeuLeuProAspThrThrGlnLysValGlnThr	196
QY	999	CGGGCTGGCGGAAACCAATCTGTTCCCCCTTATCTGGCACAGGGTGCCGCG	1058
DB	197	AspLeGlyGluAlaSerSerSer	208
QY	1059	AGGGAGACCTTTGGCCCTCTCGAGCTCGCGGTGGAGGACCTGTCTGAGGACTCGG	1118
DB	209	GlnLysProIleSerThrAspAsnAsnProAspLeuLysValAspGlyValLeuThrArg	228
QY	1119	AGCCGGAGGGCGCCACCCCG	1157
DB	229	SerGlnHisThrThrValProSerAsnIleThrIleLysLysAspGlyThrSerValGln	248
QY	1158	CGGACATTACCGCTCGCGCATGAGCCCTCCACACCGGGGGCCCAATTGTCAGCCCTC	1217
DB	249	TyrGlnHisProIleArgAsnTyr	266
QY	1218	CAGTATTGGCCCTTTCTTCGAGATCTCTATTAATTGGAATACTAACCATCCCTTC	1277
DB	267	ValArgAsnProPheArgProLeuGluLeuGlnLeuArgLysAspCysProAlaLeu	286
QY	1278	TCGGAGGATCCCAACGCTCACGGGTGTGGAGTCCCTTATGTTCTCTCACCAGCCT	1337
DB	287	ProGluGlyIleProGlnLeuAlaGluTrpLeuThrGlnThrMetAlaIleTyrAsnCys	306
QY	1338	ACTTGGGATGATTGTCACACTGTCGAGACACTTTCACACCGAGGAGCGAGAGA	1397
DB	307	AspGluAlaAspValGluGlnLeuAlaArgValIlePheProThrProValArgGln	325
QY	1398	ATTCTATTAGAGGCTAGAAAAAATGTTCTCGGGCCGACGGCGACCCACCGCGGTGCA	1457
DB	326	IleAlaGlyValIleAsnGlyHisAlaAlaAlaAsnThrAlaAlaLysIleGln	343
QY	1458	AATGAGATT	1484
DB	344	AsnTyrValThrAlaCysArgGlnHisTyrProAlaValCysAspTrpGly	360
QY	1485	ACTCGCCCGGTTCGGACTCAACACGCTGAAGGTAGGGAGGCTTGAATACTATCGC	1544
DB	361	ThrIleGlnAlaPheThrTyrLysProGlnThrAlaHisGluTyrValLysHisAla	380
QY	1545	CAGGCTCTGGTGGGGGTCTCCGGGGCGCTCAAGACGGCCCACTAATTTGGCTAAGGTA	1604
DB	381	GluIleIlePheLysAsnAsnSerGlyLeuGluTrpGlnHisAlaThrValProPheIle	400
QY	1605	AGAGAAGTATGAGGGACCGAATGAACCCCTCTGTT	1643
DB	401	AsnMetValValGlnGly	418
QY	1644	TTTCTTGAGAGGCTCTTGAAGCCTTCAGCGGTACACC	1682
DB	419	AsnProAspTrpSerThrLysThrIleProGlnIleIleProLeuMetGlnHisTyrLeu	438
QY	1683	CTTTTGTATCCACCTCAGAGGCCCAAAAAGCCCTCAGTGGCTTTGGCTTTATAGGACAG	1742
DB	439	AsnLeuGlnSerArgGlnAspAlaLysIleLysGlnThrProLeu	453
QY	1743	TCAGCCTTGATATAGAAAGACCTTCAGAGACTGGAAGGGTTACAGGAGGCTGATTA	1802
DB	454	ValLeuGlnLeuAlaMetProAlaGlnThrMetAsnGly	466
QY	1803	CGTGATCTAGTGAAGGAGGAGAGAGTATATTACAAAAGGGAGACAGAAAGAAAGG	1862
DB	467	AsnLysGlyTyrValGlySerTyrProThrAsnGluPro	479

Db 785 ArgProLeuIleSerSerLeuGluAsnGlnGlyIleLeuIleLysCysHisSerProCys 804
Qy 2796 AATACTCCCTGCTACCGTTAGAAAGCTGGGACTTAATGACTATTCAGACCGTACAGAC 2855
Db 805 AsnThrProIlePheProIleLysAlaGlyArgAspGluTyrArgMetIleHisAsp 824
Qy 2856 TTGAGAGGTCATAAAACGGGTGCAGGATATACACCAACAGTCCCGAACCCCTTATAAC 2915
Db 825 LeuArgAlaIleAsnAsnIleValAlaProLeuThrAlaValAlaIleAsnProThrThr 844
Qy 2916 CTCTTGCTGCTCTCCCAACCGAGCTGGTATACAGTATTGGACTTTAAAGGATGCC 2975
Db 845 ValLeuSerAsnLeuAlaProSerLeuHisTrpPheThrValIleAspLeuSerAsnAla 864
Qy 2976 TTCTTCTGCTGATATACCCCACTAGCCACCTCTTTTGGCTTCGATGGAGAT 3035
Db 865 PhePheSerValProIleHisLysAspSerGlnTyrLeuPheAlaPhe----- 880
Qy 3036 CCAGGTACGGGAACACCGGCGACCTCACCTGGACCGCTGCCCAAGGTTTCAAGAAC 3095
Db 881 -----ThrPheGluGlyHisGlnTyrTrpThrValLeuProGlnGlyPheIleHis 898
Qy 3096 TCCCGGACCACTTTTGACAGGCCCTACACAGACCTG-----GCCAACTTCAGGATC 3149
Db 899 SerProThrLeuPheSerGlnAlaLeuTyrGlnSerLeuHisLysIleLysPheLysIle 918
Qy 3150 CAACACCTCAGGTGACCTCTCCAGTACCTGATGACCTGCTTCTGGCGGAGCCACC 3209
Db 919 SerSer-----GluIleCysIleTyrMetAspValLeuIleAlaSerLysAsp 935
Qy 3210 AACAGGACTCTTAGAAGGACCAAGCCTACTGCTGGAAATGTCTGACCTAGGCTAC 3269
Db 936 ArgAspThrAsnLeuLysAspThrAlaValMetLeuGlnHisLeuAlaSerGluGlyHis 955
Qy 3270 AGACCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAGTAAACATATCTGGGTACAGT 3329
Db 956 LysValSerLysLysLysLeuGlnLeuCysGlnGlnGluValTyrLeuGlyGlnLeu 975
Qy 3330 TTGGCGGACGGGACGATGCTGACGAGGACGCGAAGAAACTGTAGTCCAGATCCG 3389
Db 976 LeuThrProGluGlyArgLysIleLeuProAspArgLysValThrValSerGlnPheGln 995
Qy 3390 GCCCCACCCACAGCAAAACAAATCAGACAGTCTTTTGGGACAGCTGGATTTGCGACTG 3449
Db 996 GlnProThrThrIleArgGlnIleArgAlaPheLeuGlyLeuValGlyTyrCysArgHis 1015
Qy 3450 TGGATCCCGGGGTTTGGACCTTAGACGCCCACTC---TACCCGCTAACCAAGAAAAA 3506
Db 1016 TrpIleProGluPheSerIleHisSerLysPheLeuGluLysGlnLeuLysLysAspThr 1035
Qy 3507 GGGGAA---TTCTCCTGGGCTCTGAGCAGCAGAGGAGGATTTGATGCTATCAAAAGGCC 3563
Db 1036 AlaGluProPheGlnLeuAspAspGlnGlnValGluAlaPheAsnLysLeuLysHisAla 1055
Qy 3564 CTGTGACGCGACCTGCTCGGCCCTCCCTGACCTAACTAAACCTTTTACCTTTATGTG 3623
Db 1056 IleThrThrAlaProValLeuValProAspProAlaLysProPheGlnLeuTyrThr 1075
Qy 3624 GATGAGCTAAGGAGTAGCCGGGGAGTTTAAACCCCAACCCCTAGGACCATGAGAGA 3683
Db 1076 SerHisSerGluHisAlaSerIleAlaValLeuThrGlnLysHisAlaGlyArgThrArg 1095
Qy 3684 CCTGTGCGCTACCTGTCAAGAGCTCGATCCTGTAGCAGTGGTGGCCCATATGCTGTG 3743
Db 1096 ProIleAlaPheLeuSerSerLysPheAspAlaIleGluSerGlyLeuProProCysLeu 1115
Qy 3744 AAGGCTATCGAGCTGTGGCCATCTGCTCAAGGACGCTGCAAAATTCACCTTCGGGACAG 3803
Db 1116 LysAlaCysAlaSerIleHisArgSerLeuThrGlnAlaAspSerPheIleLeuGlyAla 1135
Qy 3804 AATAAATCTGTAATAGCCCCCATGTCATTGGAGAACAATCGTTTCGGCAGCCCCCAGACCGA 3863
Db 3863

Db 1136 ProLeuIleIleTyrThrThrHisAlaIleCysThrLeuLeuGlnArgAspArgSerGln 1155
Qy 3864 TGGATGACCAACGCCCGCATGACCCACTATCAAAAGCTGTCTCTCACAGAGAGGTCACG 3923
Db 1156 LeuValThrAlaSerArgPheSerLysTrpGluAlaAspLeuLeuArgProGluLeuThr 1175
Qy 3924 TTGCTCCACAGCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGAACCA 3983
Db 1176 PheValAlaCysSerAlaValSerProAlaHisLeuTyrMetGlnSerCysGluAsnAsn 1195
Qy 3984 GTG---ACTCATGATGCCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGACCTT 4040
Db 1196 IleProProHisAspCysValLeuLeuThrHisThrIleSerArgProArgProAspLeu 1215
Qy 4041 ACAGACATACGCTGACTGGAGAAGTCTAACTGGTTCACTGACGGAACGACTATGTG 4100
Db 1216 SerAspLeuProIleProAspProAspMetThrLeuPheSerAspGlySerTyrThr 1235
Qy 4101 GTGGAAGGTAAAGAGATGGCTGGCGCGGTGTGACGGGACCCCGCACG----- 4151
Db 1236 GlyArgGly-----GlyAlaAlaValValMetHisArgProValThrAspAsp 1251
Qy 4152 ---ATCTGGGCCACGACCTGCGGAAGAACTTTCAGCAACAAAGGCTGAGCTCATGGCC 4208
Db 1252 PheIleIleIleHisGlnGlnProGlyGlyAlaSerAlaGlnThrAlaGluLeuAla 1271
Qy 4209 CTCACGCAAGCTTTGCGGCTGGCGGAGGGAATCCATAACATTTTATACGACGACGAGG 4268
Db 1272 LeuAlaAlaCysHisLysAlaThrAspLysThrValAsnIleTyrThrAspSerArg 1291
Qy 4269 TATGCCCTTTCGACCTGCACACGCTACATGCGGCCCATCTATAACAAAGGGGTTCCTTACC 4328
Db 1292 TyrAlaTyrGlyValValHisAspPheGlyHisLeuTrpMetHisArgGlyPheValThr 1311
Qy 4329 TCAGCAGGGAGGAAATAAGAACAAAGAGAAATCTTAAGCCTTATTAGAAGCCGTACAT 4388
Db 1312 SerAlaGlyThrProIleLysAsnHisLysGluIleGluTyrLeuLeuLysGlnIleMet 1331
Qy 4389 TTACCAAAAAGGCTAGCTATTATACACTGCTCTGGACATCAGAAAGCTAAAGATCTCATA 4448
Db 1332 LysProLysGlnValSerValIleLysIleGluAlaHisThrLysGlyValSerMetGlu 1351
Qy 4449 TCAGAGAGAAACACAGATGGCTGACCGGTTGCCAAGCAGGACGCCCGGGTGTAAACCTT 4508
Db 1352 ValArgGlyAsnAlaAlaAlaAspGluAlaAlaLysAsnAlaValPheLeuVal----- 1369
Qy 4509 CTGCTTAATATAGAAATGCCAAAGCCCAAGCCCAAGACCCAGACGACAGTACCTAGAGAC 4568
Db 1370 -----GlnArgValLeuLysLys 1375
Qy 4569 TGGCAAGAGATAAAAAGATAGAC-----CAGTTCTCTGAGACTCCGGAAGGGACC 4619
Db 1376 GlyAspAlaLeuAlaSerThrAspLeuValMetGluTyrSerGluThrAspGluLysPhe 1395
Qy 4620 TGCTATACCTCAGAT-----GGGAAGGAATCTCCGCCCCAC 4655
Db 1396 ThrAlaGlyAlaGluLeuHisAspGlyValPheMetArgGlyAspLeuIleValProPro 1415
Qy 4656 AAAGAAGGGTTAGAAATATGTC---CAACAGATACATCGTCTAACCCACCTA----- 4703
Db 1416 LeuGluMetLeuHisAlaIleLeuLeuAlaIleHisGlyValSerHisThrHisLysGly 1435
Qy 4704 GGAATAAACACCTGCGACGAGTGGTTCAGAACATCCCTTATCATGTTCTGAGGCTACCA 4763
Db 1436 GlyIleMetSerTyrPheSerLysPheTrpThrHisPro-----LysAlaSer 1451
Qy 4764 GGAGTGGCTACTCGTGGTCAAAATTTGTGTGCTGCCAGCTGGTGTAACTGCTTAATCCT 4823
Db 1452 GlnThrIleAspLeuIleLeuGlyHisCysGlnIleCysLeuLysHisAsnProLysTyr 1471
Qy 4824 TCCAGATGCTCCAGGAAAGAGACTAAGGGGAAGCACCAGCGGCTCATCTGGGAAGTG 4883
Db 1472 LysSerArgLeuGlnGlyHisArgProLeuProSerArgProPheAlaHisLeuGlnIle 1491

```
Qy 4884 GACTTCACTGAGTA-----AAGCCGGCTAAATACGGAACAATACTATTGTT 4934
Db      ||||| :|||:
1492 AspPheValGlnMetCysValLysPro-----MetTyrAlaLeuVal 1506
Qy 4935 TTTGTAGACACTTTTCAGGATGGGTAGAGCCTTATCCTACTAAGAAGAGACTTCAACC 4994
Db      ||||| :|||:
1507 IleIleAspValPheSerLysTrpProGluIleIleProCysAsnLysGluAspAlaLys 1526
Qy 4995 GTGGTGGCTAAATAAATACTCGAAGAAATTTTCCAGATTTCGAATACCTAAGGTAATA 5054
Db      ||||| :|||:
1527 ThrValCysAspIleLeuMetLysAspIleIleProArgTrpGlyLeuProAspGlnIle 1546
Qy 5055 GGGTCAGACAAATGTCAGCTTTTGTGCCCAGTAAGTCAGGACTGGCCCAAGATATTG 5114
Db      ||||| :|||:
1547 AspSerAspGlnGlyThrHisPheThrAlaLysIleSerGlnGluLeuThrHisSerIle 1566
Qy 5115 GGGATTGATGGAACCTGCATTGTGCATACAGACCCCAAGCTCAGACAGGTAGAGAGG 5174
Db      ||||| :|||:
1567 GlyValAlaTrpLysLeuHisCysProGlyHisProArgSerSerGlyIleValGluArg 1586
Qy 5175 ATGAATAGAACCAATAAAGAGACCTTACTAAATTGACCGGGAGACTGGGGTTAATGAT 5234
Db      ||||| :|||:
1587 ThrAsnArgThrLeuLysSerLysIleIleLysAlaGlnGluGlnLeuLysSerLys 1606
Qy 5235 TGGATAGCTCTCCCTGCTTTGTGCTTTTGGGTTAGGTTAGGAACACCCCTGGACAGTTTGGG 5294
Db      ||||| :|||:
1607 TrpThrGluValLeuProTyrValLeuLeuGluMetArgAlaThrProLysLysHisGly 1626
Qy 5295 CTGACCCCTTATGATTAATCTACTACGGGGGACCCCCCA-----TTGGTAGAAATT 5345
Db      ||||| :|||:
1627 LeuSerProHisGluIleValMetGlyArgProMetLysThrThrTyrLeuSerAspMet 1646
Qy 5346 GCTTCTGTACATAGTGTGCTGCTTTCCAGCCTTTGTTCTCTAGGCTCAAGGCA 5405
Db      ||||| :|||:
1647 SerProLeuTrpAlaThrAspThrLeuValThr----- 1657
Qy 5406 CTTGAGTGGGTGAGACAAACGCGTGGAGGCACTC-----CGGAGGCCCTAC 5453
Db      ||||| :|||:
1658 -----TyrMetAsnLysLeuThrArgGlnLeuSerAlaTyrHisGlnValVal 1674
Qy 5454 TCAGGAGGAGAGACTTGCAGATCCACATCGTTTCCAGTGGGAGATTTCAGTCTAGTT 5513
Db      ||||| :|||:
1675 AspGlnTrpProSerThrSerLeuProGlyProGluProGlySerTrpCysMetLeu 1694
Qy 5514 AGACGCCACCGTGCAGGAACCTCGAGACTCGGTGGAAGGCCCTTATCTCGTACTTTTG 5573
Db      ||||| :|||:
1695 ArgAsnProLysLys-----SerSerAsnTrpGluGlyProPheLeuIleLeuLeu 1711
Qy 5574 ACCACACCAACGGTGTGAAGTCGAAGGAATCTCCACCTGGATCCCATGCCACGTT 5633
Db      ||||| :|||:
1712 SerThrProThrAlaValLysValGluGlyArgProThrTrpIleHisLeuAspHisCys 1731
Qy 5634 AAA 5636
Db      |||||
1732 Lys 1732
```

Search completed: February 14, 2006, 16:02:52
Job time : 502.429 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 12:52:14 ; Search time 338.343 Seconds
(without alignments)
3391.447 Million cell updates/sec

Title: US-10-723-552-3

Perfect score: 14636

Sequence: 1 GCGGCGGTGCTAGCTGTGGG.....CTGTTTCATCAAAAAA 8132

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10723552/runat 14022006 125141 12876/app query.fasta_1
-DB=UniProt -PMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOFPCI=0 -LOFEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USR=US10723552 @CGN 1.1 808 @runat 14022006 125141 12876 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONLOG -DEV TIMEOUT=120
-WEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot.sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11454.5	78.3	2378	2 Q901X4_9GAMR	Q901X4 porcine end
2	11005.5	75.2	2376	2 Q901X5_9GAMR	Q901X5 porcine end
3	10997.5	75.1	2376	2 Q901X3_9GAMR	Q901X3 porcine end
4	8821	60.3	1720	2 Q8J4V6_9GAMR	Q8J4V6 porcine end
5	8815.5	60.2	1718	2 Q8J4V8_9GAMR	Q8J4V8 porcine end
6	6248.5	42.7	1193	2 Q5Q08_9GAMR	Q5Q08 porcine end
7	6170.5	42.2	1193	2 Q90RL9_9GAMR	Q90RL9 porcine end
8	6109	41.7	1195	2 Q90RL6_9GAMR	Q90RL6 porcine end
9	6026	41.2	1145	2 Q9XSN8_FIG	Q9XSN8 sus scrofa
10	5986	40.9	1145	2 Q4VFZ2_9GAMR	Q4VFZ2 porcine end
11	5863	40.1	1146	2 Q8UMP5_9GAMR	Q8UMP5 porcine end
12	5840.5	39.8	1144	2 Q8UM99_9GAMR	Q8UM99 porcine end
13	5831	39.9	1147	2 Q8UM96_9GAMR	Q8UM96 porcine end
14	5466.5	37.3	1142	2 Q8Q6U4_9GAMR	Q8Q6U4 porcine end
15	5383.5	36.8	1139	2 Q8Q6U7_9GAMR	Q8Q6U7 porcine end
16	5383.5	36.8	1736	2 Q7ZJT6_9GAMR	Q7ZJT6 amphotrophic

17	5383	36.8	1733	2 Q9E7M1_9GAMR	Q9E7M1 dg-75 murin
18	5347.5	36.5	1734	2 Q9J8E2_9GAMR	Q9J8E2 murine leuk
19	5344.5	36.5	1738	2 Q39735_MLVFR	Q39735 friend muri
20	5343.5	36.5	1734	2 Q7SVK7_9GAMR	Q7SVK7 murine leuk
21	5338	36.5	1737	2 Q92808_MLVMO	Q92808 moloney mur
22	5333.5	36.4	1738	2 Q8UN00_MLVMO	Q8UN00 moloney mur
23	5333.5	36.4	1734	2 P70355_MOUSE	P70355 mus musculu
24	5291.5	36.2	1736	2 Q5PYI3_9GAMR	Q5PYI3 murine leuk
25	5266.5	36.0	1736	2 Q83362_9GAMR	Q83362 murine leuk
26	5255	35.9	1785	2 Q89811_FLV	Q89811 murine leuk
27	5237	35.8	1784	2 Q85521_FLV	Q85521 feline leuk
28	4450	30.4	868	2 Q73505_9GAMR	Q73505 porcine end
29	4332.5	29.6	1203	2 Q89815_9GAMR	Q89815 mus dunni e
30	4316.5	29.5	1165	1 POL_GALV	P21414 gibbon ape
31	417.5	28.5	1127	2 Q9TTC1_PHACI	Q9TTC1 phascolatct
32	4154.5	28.4	1127	2 Q70652_GALV	Q70652 gibbon ape
33	3996	27.3	1196	1 POL_MLVAV	P03356 akv murine
34	3992.5	27.3	1204	1 POL_MLVFP	P26808 friend muri
35	3987.5	27.2	1199	2 Q60FS9_9GAMR	Q60FS9 murine leuk
36	3986.5	27.2	1199	2 Q60FS6_9GAMR	Q60FS6 murine leuk
37	3984.5	27.2	1204	2 Q41250_MLVRA	Q41250 rauscher mu
38	3981	27.2	1196	2 Q90RL4_9GAMR	Q90RL4 murine leuk
39	3980.5	27.2	1204	1 POL_MLVFF	P26809 friend muri
40	3977.5	27.2	1204	2 Q9YK99_9GAMR	Q9YK99 murine leuk
41	3969.5	27.1	1199	1 POL_MLVMO	P03355 moloney mur
42	3959.5	27.1	1204	1 POL_MLVFS	P26810 friend muri
43	3947	27.0	1196	1 POL_MLVRD	P11227 radiation m
44	3940.5	26.9	1189	1 POL_BAEVM	P10272 baboon endo
45	3939	26.9	1204	2 Q7ZKZ7_9GAMR	Q7ZKZ7 recombinant

ALIGNMENTS

RESULT 1
Q901X4_9GAMR
ID Q901X4_9GAMR PRELIMINARY; PRT; 2378 AA.
AC Q901X4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Type C proviral gag, pol and env genesand LTR (Class A, clone 42).
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gamma-retrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=type C;
RA Czauderna F., Fischer N., Boller K., Krach U., Kurth R., Toenjes R.R.;
RT "Molecular Characterization of Human-tropic and Replication-competent
RT Porcine Endogenous Retroviruses.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ13817; CAB65340.1; -; Genomic_DNA.
DR HSSP; P03385; 1MOF.
DR SMR; Q901X4; 7-98, 2227-2279.
DR MEROPS; A02_020;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006310; F:DNA recombination; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR GO; GO:0019068; P:RNA assembly; IEA.
DR InterPro; IPR002050; Env polyprotein.
DR InterPro; IPR008981; Env polyprotein.
DR InterPro; IPR003036; Gag p30.
DR InterPro; IPR000840; G retro matrix.
DR InterPro; IPR001995; Peptidase_A2_cat.


```

DR InterPro; IPR000223; Peptidase_S26A.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_P30; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1.
DR Pfam; PF00429; TIV_Coat; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS00501; SPASE_1; 1.
DR PROSITE; PS50158; ZF_CCHC; 1.
FT CHAIN 1 524 Gag protein.
FT CHAIN 525 1759 Pol protein.
FT CHAIN 1719 2378 Env protein.
SQ SEQUENCE 2378 AA; 267306 MW; 6EC8DA5C1B4ED18A CRC64;

Alignment Scores:
Pred. No.: 0 Length: 2378
Score: 11454.50 Matches: 2156
Percent Similarity: 94.1% Conservative: 73
Best Local Similarity: 91.0% Mismatches: 77
Query Match: 78.3% Indels: 64
DB: 2 Gaps: 8

US-10-723-552-3 (1-8132) x O9Q1X4_9GAMR (1-2378)
QY 585 ATGGGACAGCGGTGACGACCCCTCTTAGTTTGACTCTCGACCATTTGGACTGAAGTTAAA 644
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20

QY 645 TCCAGGGCTCATAATTTCTCAGTTTAAAGAGGACCTTGGCAGACTTTCTGTGTC 704
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAla 40

QY 705 TCTGAATGGCCGACATTTCGATGTGGATGGCCATCAGAGGGGACCTTTAATCTCGAGATT 764
Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60

QY 765 ATCTGGCTGTAAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 824
Db 61 IleuAlaValLysAlaIlelePheGlnThrGlyProGlySerHisProAspGlnGlu 80

QY 825 CCCTATATCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATGGGTTAAACCATGG 884
Db 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100

QY 885 CTGAATAAGCCAAAGAACCCAGGTCCCGAATTTGGCTCTTGGAGAGAAAACAAACAC 944
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120

QY 945 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCACCGGCT 1004
Db 121 SerAlaGluLysValGluProSerProArgIleTrpProGluIleGluGluProProThr 140

QY 1005 TGGCCGGAAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGGGTGCGCGAGGGGA 1064
Db 141 TrpProGluProGlnProValProProProTyrProAlaGlnGlyAlaValArgGly 160

QY 1065 CCCTTTGGCCCTCTGAGAGCTCCGGCGGTGAGGGGACCTGCTGAGGGACTCGAGCGCG 1124
Db 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180

QY 1125 AGGGGCGCCACCCGGAGCGGACGACGAGATCCGACATTTACCGCTCGCGACGTTACGGC 1184

```

```

Db 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
QY 1185 CTTCCTCCACACGGGGGCGCAATTGTCAGACCCCTCGAGATTGGCCCTTTTCTTCGCGAGAT 1244
Db 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTrpProPheSerSerAlaAsp 220
QY 1245 CTCTATATTTGGAAAACCTAACCATCCCTTCTCGGAGGATCCCAACGCGCTCACGGGG 1304
Db 221 LeuTyrAsnTrpLysThrAsnHisProPheSerGluAspProGlnArgLeuThrGly 240
QY 1305 TTGTGTGAGTCCCTTATGTTCTCTCACAGCCTACTTGGGATGATTTCAACAGCTGCTG 1364
Db 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
QY 1365 CAGACACTCTTCAACACGAGGAGGAGAGAGATTCTATTAGAGGCTAGAAAAATGTT 1424
Db 261 GlnThrLeuPheThrThrGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
QY 1425 CTTGGGGCCGACGGCGCACCCAGCGGTGCAAAATCAGATTGACATGGGATTTCCCTTA 1484
Db 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProLeu 300
QY 1485 ACTCGCCCGGTTGGGACTACACACACGCTCAAGAGTAGGAGAGCTTTCAAAATCTATCGC 1544
Db 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
QY 1545 CAGGCTCTGTGGCGGTCTCCGGGGCGCTCAAGACGGCCCTAATATTGGCTAAGGTA 1604
Db 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
QY 1605 AGAGAAGTGTAGCAGGACCGCAATGAACCCCTCTCTCTTTCTTGAGAGGCTCTTGAA 1664
Db 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
QY 1665 GCCTTCAGGGCGGTACACCCCTTTTGATCCACCTCAGAGGCCCCAAAAGCCTCAGTGGCT 1724
Db 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
QY 1725 TTGGCCCTTTATAGACAGTCCAGCCCTGGATATTAGAAGAAGCTTCAGAGCTGGAAGGG 1784
Db 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
QY 1785 TTACAGAGGCTGAGTTACGTGATCTAGTGAAGGAGCAGAGAAAGTATATTACAAGG 1844
Db 401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 420
QY 1845 GAGACAGAAGAAAGGGAACAAAGAAAAGAGAGAGAAAGAGAGAAAGGAGGAGAAAGA 1904
Db 421 GluThrGluGluGluLysGluGlnArgLysGluLysGluArgGluArgGluGluArg 440
QY 1905 CGTAAATAACCGCAAGAGAGAATTTGACTAAGATCTTGGCTGCGAGTGGTTGAAGGGAAA 1964
Db 441 ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaValAlaGluGlyLys 460
QY 1965 AGCAATACGGAACAGAGAGAGATTATTTAGGAAAATTAGTTCAGGCCCTTAGACAGTCAGGG 2024
Db 461 SerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
QY 2025 AACCTGGGCAATAGGACCCCACTCGACAGACCAATGTGCATATTGTCTAAAGAAGAGGA 2084
Db 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
QY 2085 CACTGGGCAAGGAATCTGCCCAAGAGGAAACAAAGAACCAAGGATCTTACCTAGCTTAGAA 2144
Db 501 HisTrpAlaArgAsnCysProLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
QY 2145 GAACATAAAGATTAGGAGGAGACGGGTTTCGACCCCTTCGACCGCCAGCGGTAACTTTG 2204
Db 521 GluAspLysAsp---GlyArgArgGlySerAspProLeuProGluProArgValThrLeu 539
QY 2205 AAGGTGAGGGGCGCAACCAAGTTGAGTTCTGTTGATATACCGGAGCGAAACATTCAGTGTGA 2264
Db 540 LysValGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeu 559

```

QY 2265 CTACAGCCATTAGGAAACTAAAGATATAAAATCTCTGGGTGATGGGTGCCACAGGGCAA 2324
DB LeuGlnProLeuGlyLysLeuLysGluLysSerTrpValMetGlyAlaThrGlyGln 579
QY 2325 CAACAGTATCCATGGACTACCCGAGAACACAGTTGACTTGGAGTGGGACGGGTAAACCCAC 2384
DB ArgGlnValArgGluPheLeuGlyThrArgValAspLeuGlyValGlyArgValThrHis 599
QY 2385 TCGTTTCTGGTCTATACCTGAGTGCACACCCCTCTTAGTGAGAGACTTATTAGCAACAG 2444
DB SerPheLeuValIleProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLys 619
QY 2445 ATGGAGCACAAATTTCTTTTGAACAAGGAAACACAGAGTGTCTGCAATATAACAACCT 2504
DB MetGlyAlaGlnIleSerPheGluGlnGlyArgProGluValSerValAsnAsnLysPro 639
QY 2505 ATCACTGTGTGACCTCCCAATTAGATGACGAATATCGACTATCTCTCCCTAGTAAAG 2564
DB IleThrValLeuThrLeuGlnLeuAspGluTyrArgLeuTyrSerProGlnValLys 659
QY 2565 CCTGATCAAAATATACAAATCTGGTTGGAACAGTTTCCCAAGCCTGGGCAGAAACCGCA 2624
DB ProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAla 679
QY 2625 GGGATGGTTTGGCAAGCAAGTTCCCCCAAGATTATCACTGAAGGCCAGTGCACCA 2684
DB GlyMetGlyLeuAlaLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThr 699
QY 2685 CCAGTCTCAGTCAGACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCGCAT 2744
DB ProValSerValArgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHis 719
QY 2745 GTCCAAAGATTAATCAACAGGGCATCTAGTTCTCTGTCCAATCTCCCTGGAATCTCC 2804
DB ValGlnArgLeuIleGlnGlyIleLeuValProValGlnSerProTrpAsnThrPro 739
QY 2805 CTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCTACAGGACTTTGAGAGAG 2864
DB LeuLeuProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGlu 759
QY 2865 GTCAATAACGGGTGAGGATATACACCCGAAGTCCCGAACCTTTATACCTCTTGTGT 2924
DB ValAsnLysArgValGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSer 779
QY 2925 GCTCTCCCAACCCAGAGCTGGTATACAGTATTGGACTTAAAGGATGCTTCTTCTGC 2984
DB AlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPheCys 799
QY 2985 CTGAGATTACCCCACTAGCCCAACCACTTTTTCCTTGAATGGAGATCCAGGTACG 3044
DB LeuArgLeuHisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThr 819
QY 3045 GGAAGAACCGGCGAGCTCAGTGGACCCGACTCCCGAGGTTCAAGACTCCCGACC 3104
DB GlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThr 839
QY 3105 ATCTTTGACCAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCAAACCCCTCAGGTG 3164
DB IlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheLysIleGlnHisProGlnVal 859
QY 3165 ACCCTCTCTCAGTACGTGGTACCTGCTTCTGGCGGAGCCCAACCAAGGACTGCTTA 3224
DB ThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeu 879
QY 3225 GAAGGCACGAGGCACTACTGTGGAAATTTCTCGACTAGGCTACAGACCTCTGTAAG 3284
DB GluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLys 899
QY 3285 AGGGCCCAATTTGACGAGAGAGGTAAACATCTTGGGGTACAGTTTGGGACGGGCGAG 3344
DB LysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGln 919

QY 3345 CGATGGCTGACGAGGACACGAAAGAAACATGTAGTCCAGATACCGGCCCCAACACACAGCC 3404
DB ArgTrpLeuThrGluAlaArgLysLysThrValValGlnIleProAlaProThrThrAla 939
QY 3405 AAACAAATGAGAGAGTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTT 3464
DB LysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPhe 959
QY 3465 GCACACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAAAGGGGAATTTCTCTCGGCT 3524
DB AlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGlyGluPheSerTrpAla 979
QY 3525 CTGAGCACACAGAGGATTTGATGTATCAAAAAGCCCTGTGTAGCGCACCTGCTCTG 3584
DB ProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeu 999
QY 3585 GCCTCCCTGACGCTAAACCCCTTACCTTTATGTGGATGAGCTGAGGAGGTAGCC 3644
DB AlaLeuProAspValThrLysProPheThrLeuTyrValAspGluArgLysGlyValAla 1019
QY 3645 CGGGAGATTTTAAACCCAAACCTTAGGACCATGAGAGACCTCTCGCTACCTGTCAAAG 3704
DB ArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgProValAlaLysLeuSerLys 1039
QY 3705 AAGCTCGATCTGTAGCAGTGGTTGGCCCATATGCTGAAGCTATCGCAGCTGTGCC 3764
DB LysLeuAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaValAla 1059
QY 3765 ATACTGGTCAGGACGCTGACAAATTTGACTTTGGGACAGATATTAATGTAAGCCCC 3824
DB IleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaPro 1079
QY 3825 CATGCATTGAGAAACATCGTTCCGCGACCCCAAGCCGATGGATGACCAACGCCCGCATG 3884
DB HisAlaLeuGluAsnIleValArgGlnProProAspArgTrpMetThrAsnAlaArgMet 1099
QY 3885 ACCACTATCAAGCCTGCTTCTCACAGAGAGGTCACGTTCCCTCCACAGCCGCTCTC 3944
DB ThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeu 1119
QY 3945 AACCTCGCACTCTTCTGCTGAGAGACTGATGAACCACTGACTCATGATGCCATCAA 4004
DB AsnProAlaThrLeuLeuProGluGluThrAspGluProValThrHisAspCysHisGln 1139
QY 4005 CTATTGATTGAGGAGACTGGGGTCCCAAGAGCCTTTACAGACATACCGCTGACTGGAGAA 4064
DB LeuLeuIleGluThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGlu 1159
QY 4065 GTGCTAACTGTGTTCACTACCGAAGCAGCTATGTGCTGGAAGGTAAAGAGATGGCTGG 4124
DB ValLeuThrTrpPheThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGly 1179
QY 4125 GGGCGGTGGTGGACGGGACCGCAGATCTGGCCAGCAGCTGCGGAGAGAACTTCA 4184
DB AlaAlaValValAspGlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSer 1199
QY 4185 GCACAAAAGGCTGAGCTCATGGCCCTCAGCAAGCTTTGCGGCTGGCCGAGGAAATCC 4244
DB AlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSer 1219
QY 4245 ATAAACATTATACGACAGCAGGTATGCTTTTGGACTGCACACGTACATGCGGGCATC 4304
DB IleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIle 1239
QY 4305 TATAACAAAGGGGTGTTGCTTACCTCAGCAGGAGGGAATTAAGAACAAAGAGGAATTT 4364
DB TyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIle 1259
QY 4365 CTAAGCCTATTAGAGCCCTACATTTTACCAAAAGGCTAGCTATTATATACCTGCTCTGGA 4424
DB LeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGly 1279
QY 4425 CATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATGGCTGACCGGGTTGCCAAG 4484

Db 1280 HisGlnLysAlaLysAspLeuLeuSerArgGlyAsnGlnMetAlaAspArgValAlaLys 1299
Qy 4485 CAGCAGCCCGGGTGTAACTTCCTGCTATATAGAAATGCCAAAGCCCGAGAACCC 4544
Db 1300 GlnAlaAlaGlnAlaValAsnLeuLeuProIleGluThrProLysAlaProGluPro 1319
Qy 4545 AGACGACGTACACCCCTAGAGACTGGCAAGATATAAAGATAGACCGATTCCTCTGAG 4604
Db 1320 ArgArgGlnThrLeuGluAspTrpGlnGluLeuLysIleAspGlnPheSerGlu 1339
Qy 4605 ACTCCGAAGGAGCCTGCTATACCTCAGATGGGAAGAAATCTCCGCCCAAGAAAGGG 4664
Db 1340 ThrProGluGlyThrCysTyThrSerTyrgLysGluLeuLeuProHisLysGluGly 1359
Qy 4665 TTAGAATATGTCACACAGATACATCGTCTAACCCACCTAGAGAACTAAACACCTGCAGCAG 4724
Db 1360 LeuGluTyValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGln 1379
Qy 4725 TTGGTCAAGACATCCCTTATCATGTCTGAGGCTACCGAGGTGGCTGCTGGTGGTC 4784
Db 1380 LeuValArgThrSerProTyHisValLeuArgLeuProGlyValAlaAspSerValVal 1399
Qy 4785 AAACATTGTGTGCCCTGCCAGCTGTTAAATGCTTAATCTTCCAGAATGCCCTCCAGGGAAG 4844
Db 1400 LysHisCysValProCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLys 1419
Qy 4845 AGACTAAGGGAAACCCACCGCGCTCAGTGGGAAGTGAATTCATCTGAGGTAAAGCCG 4904
Db 1420 ArgLeuArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluValLysPro 1439
Qy 4905 GCTAAATACGGAACAAATACCTATTGTTTGTAGCACCTTTCCAGGATGGGTAGAG 4964
Db 1440 AlaLysTyrgLysAsnLysTyLeuLeuValPheValAspThrPheSerGlyTrpValGlu 1459
Qy 4965 GCTTATCTCTAAGAAAGAGACTTCAACCGTGTGGTGAATAAATACTCGAAGAAAT 5024
Db 1460 AlaTyProThrLysLysGluThrSerThrValValAlaLysLysIleLeuGluGluLe 1479
Qy 5025 TTTCACAGATTGGAATCCTAAGGTAAATAGGTCAGACAAATGTCAGCTTTTGTGCC 5084
Db 1480 PheProArgPheGlyIleProLysValIleGlySerAspAsnGlyProAlaPheValAla 1499
Qy 5085 CAGTGAAGTCAGGACCTGGCCAGATATTGGGATTCATTTGGAACTGGCAATGTCATAC 5144
Db 1500 GlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTy 1519
Qy 5145 AGACCCCAAGCTCAGGACAGGTAGAGGATGAATAGAACCATTAAGAGACCTTACT 5204
Db 1520 ArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThr 1539
Qy 5205 AAATTGACCGGGAGACTGGCGTTAATGATTTGATAGTCTCTCCCTTGTGCTTTT 5264
Db 1540 LysLeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPhe 1559
Qy 5265 AGGTTTAGGACACCCCTGGACACTTTGGGCTGACCCCTATGATTAATCTACTACGGGGA 5324
Db 1560 ArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrgLysLeuLeuLeuTyrgLys 1579
Qy 5325 CCCCCCATTTGGTAGAAATGCTTCTGTACATAGTGTGACGTGCTTCCAGCTTTCCAGCCT 5384
Db 1580 ProProProLeuValGluIleAlaSerValHisSerAlaAspValLeuLeuSerGlnPro 1599
Qy 5385 TTGTTCTTAGGCTCAAGGCACTTGAGTGGGTAGACAACGAGCGTGGAGCAACTCCGG 5444
Db 1600 LeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArg 1619
Qy 5445 GAGGCTTACTCAGGAGGAGGAGCTTGCAGATCCCATCGTTTCCAAGTGGGAGATTCA 5504
Db 1620 GluAlaTySerGlyGluGlyAspLeuGlnValProHisArgPheGlnValGlyAspSer 1639
Qy 5505 GTCTACGTTAGACGCCACCGTGCAGGAACCTCCAGACTCGGTGGAGGGCCCTTATCTC 5564

Db 1640 ValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyLeu 1659
Qy 5565 GTACTTTTGACCAACACCAACCGCTGCAAGTTCGAAGATCTCCACCTGCATCCATGCA 5624
Db 1660 ValLeuLeuThrThrProThrAlaValLysValGluArgIleProThrTrpIleHisAla 1679
Qy 5625 TCCACACTTAAACCGCGCCACCTCCCGATTCGGGGTGGAAAGCCGAAAAAGACTGAAAAAT 5684
Db 1680 SerHisValLysProAlaProProProAspSerGlyTrpLysAlaGluLysThrGluAsn 1699
Qy 5685 CCCCTTAAGCTTCGCCCTCCATCGGTGTTCTTACTCTCTCAATAAACCTCT 5736
Db 1700 ProLeuLysLeuArgLeuHisArgValValProTySerValAsnAsnSerSerSerMe 1719
Qy 5736 ----- 5736
Db 1719 tHisProThrLeuSerArgArgHisLeuProIleArgGlyLysProLysArgLeuLys 1739
Qy 5737 -----CAGAC 5741
Db 1739 sIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGlnVa 1759
Qy 5742 TAATGTTATCGCATAGGACAGACCTGAACTCCATAAAACCTTATCTCTCACCTGGTT 5801
Db 1759 lAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrpLe 1779
Qy 5802 AATTACTGACTCCGGCACAGGTATTATATCAACACACTCAAGGGGAGGCTCCTTAGG 5861
Db 1779 uLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeuGl 1799
Qy 5862 AACCTGTGTGCTGCTATATCTATCTGCTCAGATCAGTTATTCTCTAGTCTCTG 5913
Db 1799 yThrTrpTrpProGluLeuTyValCysLeuArgSerValIleProGlyLeuAsnAspGl 1819
Qy 5914 -ACCTCAACCCGATATCTCCATCTCAGGATTTATGTTGTCAGGACACCAAA 5972
Db 1819 nAlaThrProProAspValLeuArgAlaTyrgLysPheTyValCysProGlyProProAs 1839
Qy 5973 TAATGGAACATTTGCGAAATCCAGAGATTTCTTTGTAAACAAATGGAAGTGTGAAC 6032
Db 1839 nAsnGluGluTyrgCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysIleTh 1859
Qy 6033 CTCTAATGATGATATTGGAATGCGCAACTCTCAGCAGAGTAGGTTAGTTTCTTCTTA 6092
Db 1859 rSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTySerPh 1879
Qy 6093 TGTCAACACCTATACCACTCTCGACCAATTAATTAC-----CTGACCTG 6137
Db 1879 eValAsnAsnProThrSerTyrgAsnGlnPheAsnTyrgLysHisGlyArgTrpLysAspTr 1899
Qy 6138 G-----ATTAGAACTGGNAGCCCAAGTCTCTCTTCAGACCT 6176
Db 1899 pGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAspLe 1919
Qy 6177 AGATTACTTAAATAAGTTTCTCTGAGAAAGGAAAAACAAGAAATATCTTAAATGGGT 6236
Db 1919 uAspTyLeuLysIleSerPheThrGluLysGlyLysGlnGlnLysIleGlnLysTrpVa 1939
Qy 6237 AAATGGTATCTCTGGGAAATGGTATATTATGAGGCTCGGGTAAACAACAACAGGCTCAT 6296
Db 1939 lAsnGlyIleSerTrpGlyIleValTyrgLysGlySerGlyArgLysLysGlySerVa 1959
Qy 6297 TCTAACTATTTCGCTCAAAATA---AACCACTGGAGCTCCATGCTATGAGCCACAAA 6353
Db 1959 lLeuThrIleArgLeuArgIleGluThrGlnMetGluProValAlaIleGlyProAs 1979
Qy 6354 TACGCTCTTCAGCGGTCAAAAGACCCCAACCCAA-----GGACGAGCACCATC 6401
Db 1979 nLysGlyLeuAlaGluGlnGlyProProIleGlnGlnArgProSerProAsnProSe 1999
Qy 6402 CTCT---AACATACTTCTGGATCAGACCCACCTGAGCTTAACAGCACGACTAAATGGG 6458
Db 1999 xAspTyAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThrGl 2019

```
QY 6459 GGCAGAACTTTTAGCCCTCATCCAGGAGAGCTTTTCAAGCTCTTAACTCCAGCTCCAGA 6518
DB 2019 YAlaLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAenSerThrThrProG 2039
QY 6519 GGCTACCTCTCTTGTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGAATGCC 6578
DB 2039 uAlaThrSerSerCysIlePheCysLeuAlaSerGlyProProIleThrGluGlyMetAl 2059
QY 6579 TAGAAGAGGGAATTCATGTGACAAAGAACATAGAGACCAATGCACATGGGGATCCCA 6638
DB 2059 aArgGlyGlyPheAenValThrLysGluHisArgAspGlnCysThrIlePheGlySerG 2079
QY 6639 AATAAGCTTACCTTACTGAGTTTCTGGAAAGGCACCTGCATAGGAAGGTTCCTCCC 6698
DB 2079 nAenLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValProP 2099
QY 6699 ATCCACCAACACCTTTGTACCACTGAAAGCTTTAATCAACCTCTGAGAGTCATA 6758
DB 2099 oSerHisGlnHisLeuCysAenHisThrGluAlaPheAenArgThrSerGluSerGlnTy 2119
QY 6759 TCTGGTACCTGGTTATCACAGGTGGTGGGCATGTAACTACCTGGATTACCCCTGTGTTC 6818
DB 2119 rLeuValProGlyTyraAspArgIleThrAlaCysAenThrGlyLeuThrProCysValSe 2139
QY 6819 CACCTTGGTTTAAACCAACTAAAGATTTTGCATTATCGTCCAAATTCCTCCCGAGT 6878
DB 2139 rThrLeuValPheAenGlnThrLysAspPheCysValMetValGlnIleValProArgVa 2159
QY 6879 GTATTACTATCCGAAAGCAATCTTGTGATGATATGATACAGAAATCATCGACAAA 6938
DB 2159 lTyTyTyTyProGluLysAlaValLeuAspLysTyraAspTyraArgTyraArgProLy 2179
QY 6939 GAGAGAACCATATCTCGACACTGTGTGATGCTCGAGTGGAGTGGCAGCAGGTGT 6998
DB 2179 sArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVa 2199
QY 6999 AGGAACAGGAACAGTGCCTGTGTACGGACCAACAGCAGCTAGAACAGGACTTAGTA 7058
DB 2199 lGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSerAs 2219
QY 7059 CCTACATCGAATGTACAGAACTCTCCAGCCCTAGAAATCTGTACGTAACCTGGA 7118
DB 2219 nLeuHisArgGlyValThrGluAspLeuGlnAlaLeuGluLysSerValSerAenLeuG 2239
QY 7119 GGAATCCCTAACCTCTTATCTGAAGTAGTCCTACAGAAATAGAGAGGGTTAGATTATT 7178
DB 2239 uGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLe 2259
QY 7179 ATTTCTAAAGAGGAGGATTTATGTAGCTTTGAAGAGGAAATGCTGTTTTTATGTGA 7238
DB 2259 uPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyValAs 2279
QY 7239 TCATTACGGGCCCATCAGACACTCCATGAACAAGCTTAGAAAGGTGGAGAAAGCGTCG 7298
DB 2279 pHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 2299
QY 7299 AAGGGAAGGAAACTACTCAAGGGTGGTTGAGGATGGTTCACAGGCTCTCTTGGTT 7358
DB 2299 sArgGluArgGluAlaAspGlnGlyThrPheGluGlyThrPheAsnArgSerProIleP 2319
QY 7359 GGCTACCTCTTCTTCTTAAACAGGATTAATGCTTCATTAGAACGAATAGTGCAGTCCA 7418
DB 2319 tThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThrVa 2339
QY 7419 TGGGCCATGATTATTAAACAGTAAATGCTTCATTAGAACGAATAGTGCAGTCCA 7478
DB 2339 lGlyProCysLeuIleAenArgPheValAlaPheValArgGluArgValSerAlaValG 2359
QY 7479 GATCATGGTACTTAGACAAACAGTACCAG 7506
DB 2359 nIleMetValLeuArgGlnGlnIleThrGln 2369
```

```
RESULT 2
Q9Q1X5_9GAMR
ID Q9Q1X5_9GAMR PRELIMINARY; PRT; 2376 AA.
AC Q9Q1X5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Type C proviral gag, pol and env genes and LTR (class B, clone
DE 33).
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammaretrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Type C;
RX MEDLINE=20219394; PubMed=10756014;
RX DOI=10.1128/JVI.74.9.4028-4038.2000;
RA Czauderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;
RT "Establishment and characterization of molecular clones of porcine
RT endogenous retroviruses replicating on human cells.";
RL J. Virol. 74:4028-4038(2000).
DR ENBL; AJ133816; CAB65339.1; -; Genomic_DNA.
DR HSP; Q9Q1X5; 7-98, 472-518, 2225-2277.
DR MEROPS; A02.020; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0004190; F:Aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR002050; Env_Polyprotein.
DR InterPro; IPR008981; FmLVrecept-bind.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001985; Peptidase_A2_cat.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF00429; TLV_coat; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00879; RNASE_H; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.
FT CHAIN 1 524 Gag protein.
FT CHAIN 525 1761 Pol protein.
FT CHAIN 1720 2376 Env protein.
SQ SEQUENCE 2376 AA; 266625 MW; 37DD39D3099DA4AF CRC64;

Alignment Scores:
Pred. No.: 0 Length: 2376
Score: 11005.50 Matches: 2078
Percent Similarity: 91.5% Conservative: 99
Best Local Similarity: 87.3% Mismatches: 118
Query Match: 75.2% Indels: 84
DB: 2 Gaps: 12
```

[illegible]

QY 2745 GTCCAAAGATTAAATCCAAAGGGGCAATCTAGTTCCTGTCCTCAATCTCCTCGGAATACTCCC 2804
Db ValGlnArgLeuIleGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrPro 739
QY 2805 CTGCTACCGGTTAGAAAGCCCTGGCACTAAATGACTATGACCACTACAGGACTTCAGAGAG 2864
Db LeuLeuProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGlu 759
QY 2865 GTCAATAAAGCGGTGCGAGATATACACCACACAGTCCCGAACCTTATAACCTCTCTGT 2924
Db ValAsnLysArgValGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSer 779
QY 2925 GCTCTCCACCCAAACGAGCTGTGTATACAGTATTGGACTTAAAGATGCTTCTCTGC 2984
Db AlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCys 799
QY 2985 CTGAGATTACACCCCACTAGCAACCACTTTTCCCTTCCGAATCGAGAGATCCAGGTACG 3044
Db LeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlnTrpArgProGlyThr 819
QY 3045 GGAAGAACCGGCGCACTCACCTGACCCGACTGCCCCCAAGGGTTCAAGAACTCCCGCAC 3104
Db GlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThr 839
QY 3105 ATCTTGACGAGCCCTACACAGAGACTGGCCCACTTCAGATCCCAACACCTCTCAGGTG 3164
Db IlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnVal 859
QY 3165 ACCTCTCTCAGTACGTGGATGACCTGCTTCTGGCGGAGCCCAACACAGGACTGTTA 3224
Db ThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThrLysGlnAspCysLeu 879
QY 3225 GAAGGCACGAAGGCACCTACTGCTGGAATTGTCTGACCTAGCTACAGAGCTCTGCTAAG 3284
Db GluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLys 899
QY 3285 AAGCCCAAGATTTCAGAGAGAGGTAACTACTTGGGTACAGTTTGGCGGACGGGCAG 3344
Db LysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGln 919
QY 3345 CGATGGCTGACGGGACCGAAGAACTGTAGTCCAGATACGGGCCCCCAACACAGCC 3404
Db ArgTrpLeuThrGluAlaArgLysLysThrValValGlnIleProAlaProThrThrAla 939
QY 3405 AAAACAAATGACAGATTTTGGGACAGCTGGATTTTGACAGCTGTGGATCCCGGGTTT 3464
Db LysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPhe 959
QY 3465 GCGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAAAGGGGAATTCTCTGGGCT 3524
Db AlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGlyGluPheSerTrpAla 979
QY 3525 CCTGAGCACAGAGGATTTGATGCTATCAAAAGCCCTGTGAGCGGCACCTGCTCTG 3584
Db ProGluHisGlnLysAlaPheAspAlaIleLysAlaLeuLeuSerAlaProAlaLeu 999
QY 3585 GCCTCTCCTCACGTAACTAAACCTTTTACCCTTTATGTGATGAGCTAAGGAGTAGCC 3644
Db AlaLeuProAspValThrLysProPheThrLeuTyrValAspGluArgLysGlyValAla 1019
QY 3645 CGGGAGTTTAAACCCAAACCTTAGGACCATGAGAGACCTGTGCGCTTACTGTCAAG 3704
Db ArgGlyValLeuThrGlnThrLeuGlyProTrpArgProValAlaTyrLeuSerLys 1039
QY 3705 AAGCTGATCTGTAGCAGTGGTTGGCCCATATGCTGAAGCTATTCGAGCTGTGGCC 3764
Db LysLeuAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaValAla 1059
QY 3765 ATACTGTTCAAGACGCTGACAAATTGACTTTGGGACAGATATACTGTAAATAGCCCC 3824
Db IleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaPro 1079
QY 3825 CATGCATTGGAAACATCGTTTCGGCAGCCCCCAGACCGATGGATGACCAACGCCCGCATG 3884

Db HisAlaLeuGluAsnIleValArgGlnProProAspArgTrpMetThrAsnAlaArgMet 1099
QY 3885 ACCCACTATCAAAAGCCCTGCTTCTCACAGAGAGGTTCAGTTCGCTCCACAGCCGCTCTC 3944
Db ThrHisTyrGlnSerLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeu 1119
QY 3945 AACCTCCCACTCTTCTGCTGAAGACAGCTGATGAACAGTCACTCATGATTGCCATCAA 4004
Db AsnProAlaThrLeuLeuProGluGluThrAspGluProValThrHisAspCysHisGln 1139
QY 4005 CTATTGATTGAGAGACTGGGGTCCGCAAGAACCTTACAGACATACACGCTGACTGGAGAA 4064
Db LeuLeuIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGlu 1159
QY 4065 GTGCTAACCTGGTTCACTGACGGAAGCAGCTATGTGGTGAAGTAAGAGATCGCTGGG 4124
Db ValLeuThrTrpPheThrAspGlySerTyrValValGluGlyLysArgMetAlaGly 1179
QY 4125 GCGCGGTGTGGACGGGACCCGACGATCTGGGCCACAGACGCTGCCGGAAGGAACCTTCA 4184
Db AlaAlaValValAspGlyThrHisThrIleTrpAlaSerSerLeuProGluGlyThrSer 1199
QY 4185 GCACAAAGGCTGAGCTCATGGCCCTCACGCAAGCTTTTGGCGTGGCCGAAGGGAATCC 4244
Db AlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSer 1219
QY 4245 ATAAACATTTATACGACAGCAGCTATGCTTCCGACTGCACACGCTACATCGGCGCATC 4304
Db IleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIle 1239
QY 4305 TATAACAAAGGGGTTGCTTACTCTCAGCAGGAGGGGAAATAAGAACAAAGAGAGAAAT 4364
Db TyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysLeuLysGluGluIle 1259
QY 4365 CTAAGCCTATTAGAAGCCGTACATTTACCAAAAGGCTAGCTATTATACACTGTCTGGA 4424
Db LeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGly 1279
QY 4425 CATCAGAAAGCTAAAGTCTCATATCCAGAGAAACCCAGATGGCTGACCGGTGCCAAG 4484
Db HisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLys 1299
QY 4485 CAGCGACCCAGGGTGTAACTTCTGCTCTATATAGAAATGCCCAAGCCCCAGAACCC 4544
Db GlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThrProLysAlaProGluPro 1319
QY 4545 AGACGACGTACACCTTAGAAGACTGGCAAGAGATAAAAAAGATAGACCAAGTTCTCTGAG 4604
Db ArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGlu 1339
QY 4605 ACTCCGGAAGGGACCTGCTATACCTCAGATGGGAAGGAATCTGCCCCACCAAGAGGG 4664
Db ThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGly 1359
QY 4665 TTAGAATATGTCACACAGATACATCGTCTAACCCCACTAGAACTAAACACCTGCACAG 4724
Db LeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGln 1379
QY 4725 TTGGTCAGAACATCCCTTATCATGTTCTGAGCTTACAGGAGTGGCTGACTCGGTGGTC 4784
Db LeuValArgThrSerProTyrHisValLeuArgLeuProGlyValAlaAspSerValVal 1399
QY 4785 AAACATTGTGTGCCCTGCCAGCTGGTGAATGCTTAATCTTCCAGATGCTCCAGGGAAG 4844
Db LysHisCysValProCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLys 1419
QY 4845 AGACTAAGGGGAGGCCACCCAGCGCTCACTGGGAAGTGGACTTCACTAGGTAAAGCCG 4904
Db ArgLeuArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluValLysPro 1439
QY 4905 GCTAAATACGGAAACAAATACTATTGGTTTTTTGTAGACACCTTTTTCAGGATCGGTAGAG 4964

Db 1440 AlalysTyrglyAsnLysTyLeuLeuValPheValAspThrPheSerGlyTrpValGlu 1459
Qy 4965 GCTTATCTACTAAGAAGACATTCACCGTGTGGCTGCTAAATAAATACTGGAGAAATT 5024
Db 1460 AlaTyProThrLysGluThrSerThrValValAlaLysLysIleLeuGluGluIle 1479
Qy 5025 TTTCGAAGATTGGAAATACCTAAGGTAAATAGGTCAGACAAATGGTCAGCTTTTGTGGCC 5084
Db 1480 PheProArgPheGlyIleProLysValIleGlySerAspAsnGlyProAlaPheValAla 1499
Qy 5085 CAGTAAAGTCAGGACCTGGCCAGCATATTGGGATTCATTGGAACTGCATTGTCATAC 5144
Db 1500 GlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTy 1519
Qy 5145 AGACCCAAAGCTCAGACAGGTAGAGAGATGAATAGAACCATTAAGAGACCTTACT 5204
Db 1520 ArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThr 1539
Qy 5205 AAATTGACCGGAGACTGGCGTTAAATGATTGGATAGCTCTCCCTGCTTGTCTTTT 5264
Db 1540 LysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeuLeuProPheValLeuPhe 1559
Qy 5265 AGGTTAGGAACACCCCTGGACAGCTTTGGGCTGACCCCTATGAATTACTCTACGGGGGA 5324
Db 1560 ArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrgluLeuLeuTyrgly 1579
Qy 5325 CCCCCCGATTGGTAGAAATTGCTTCGTACATAGTCTGACGTGCTTCCAGCCT 5384
Db 1580 ProProLeuValGluIleAlaSerValHisSerAlaAspValLeuLeuSerGlnPro 1599
Qy 5385 TTGTTCTCTAGGCTCAAGGCACTTGAAGTGGGTGAGACACGAGCGTGAGGCACTCCGG 5444
Db 1600 LeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArg 1619
Qy 5445 GAGGCTACTCAGGAGGAGAGACTTCAGATCCCACATCGTTTCCAAGTGGGAGATTCA 5504
Db 1620 GluAlaTySerGlyGluGlyAspLeuGlnValProHisArgPheGlnValGlyAspSer 1639
Qy 5505 GTCTACGTTAGACGCCACCGTGAGGAACCTCAGACTCGGTGAGAGGCCCTTATCTC 5564
Db 1640 ValTyValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrgly 1659
Qy 5565 GTACTTTTGACACACCAACCGCTGCAAGTCAAGGAATCTCCACTGGATCCATGCA 5624
Db 1660 ValLeuLeuThrThrProThrAlaValLysValGluGlyIleSerThrTrpIleHisThr 1679
Qy 5625 TCCACGTTAAACCGCGCCACCTCCCGATTGGGGTGGAAAGCCGAAAGACTGAAAT 5684
Db 1680 SerHisValLysLeuAlaProProAspSerGlyTrpArgAlaGluLysThrGluAsn 1699
Qy 5685 CCCCTTAAGCTTCGCCCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCT 5736
Db 1700 ProLeuLysLeuArgLeuHisArgLeuValProTyrglySerAsnAsnAsn-SerProGly 1719
Qy 5736 ----- 5736
Db 1719 nIleHisProThrLeuSerTrpArgHisLeuLeuProThrArgGlyGluProLysArgLe 1739
Qy 5737 -----CA 5738
Db 1739 waTGileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleThrProG 1759
Qy 5739 GACTAATGGTATGGCATAGAGACAGCTGAACTCCGATAAACCTTATCTTCACGCTG 5798
Db 1759 nAlaSerSerLysArgLeuIleAspSerSerAsnProHisArgProLeuSerLeuThrTr 1779
Qy 5799 GTTAATTACTCACCGGCACAGGTATTAAATATCAACAACTCAAGGGGAGGCTCCTTT 5858
Db 1779 pLeuIleLeuAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProAr 1799
Qy 5859 AGGAACCTGGTGGCTCATCTATACGTTGCTCGCATCAGTATTCTAGCTGACCTC 5918
Db 1799 gGlyThrTrpTrpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSe 1819

Qy 5919 A----CCCCAGATATCTCATCTACGGATTATTATGTTTGCACGAGCCACCAATAA 5975
Db 1819 rThrProProAsnLeuValArgSerTyrglyPheTyrglyCysProGlyThrGluLysG 1839
Qy 5976 TGGAAACATTTGGGAAATCCACAGAGATTTCTTTGTAAACAATGGAACCTGTGTAACCTC 6035
Db 1839 u---LysTy-CysGlyAspSerGlyGluSerPheCysArgArgTrpSerCysValThrSe 1858
Qy 6036 TAATGATGGATATTGGAAATGGCCAACTCTCAGCAGGATAGGTAAGTTTCTTATGT 6095
Db 1858 rAsnAspGlyAspTrpLysTrpIleSerLeuGlnAspArgValLysPheSerPheVa 1878
Qy 6096 CAACACCTATACCAGCTCTCGACAAATTTAACTTACCTGACCTGGATTAGAACTGGAAGCC 6155
Db 1878 lAsn-----SerGlyProGlyLysTyrglyValMetLysLeuTyrglyAspLy 1894
Qy 6156 CAAGTGTCTCTTTCAGACCTAGATTACCTAAAAATAAGTTTCACTGAGAAAGAAACA 6215
Db 1894 sSerCysSerProSerAspLeuAspTyrglyLysIleSerPheThrGluLysGlyLysG 1914
Qy 6216 AGAAATATCTTAAATGGTAAATGGTATGCTCTTGGGGAATGGTATAT-----TATGG 6269
Db 1914 nGluAsnIleGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrglyTrG 1934
Qy 6270 AGGCTCGGTAAACAACCGCTCCATTCTTAATATTTCGCTCAAAATAAACCCAG---CT 6326
Db 1934 yGlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyTh 1951
Qy 6327 GAGCCTCCAAATGCTATAGGACCAATACGCTCTTGACGGGTCAAAAGACCCCA----- 6381
Db 1951 rGluProProValAlaValGlyProAspLysValLeuAlaGluGlnGlyProAlaLe 1971
Qy 6382 -----ACCA 6386
Db 1971 uGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrG 1991
Qy 6387 AGGACGAGGACCTCTCTAACATACTCTCGATCAGACCCCACTCAGTCT-----AA 6440
Db 1991 nSerPro-----SerAsnGlyThrThrGlyLeuIleProThrAsnThrProArgAs 2008
Qy 6441 CAGC-----ACGACTAAATGGGGCAAACTTTTAGCCTCATCCAGGAGCTTT 6491
Db 2008 nSerProGlyValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPh 2028
Qy 6492 TCAAGCTCTTAACCTCAGACTCCAGAGGCTACCTCTTCTTGTGGTATGCTTAGCTTC 6551
Db 2028 eGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSe 2048
Qy 6552 GGGCCACCTTACTATCAAGGAATGGCTAGAAGAGGGAATTCATGTGCAAAAGAACA 6611
Db 2048 rGlyProProTyrglyGluGlyMetAlaLysGluGlyAspPheAsnValThrLysGluHi 2068
Qy 6612 TAGAGACCAATGACATGGGGATCCCAAAATAAGCTTTACCTTACTGAGGTTCTGGA 6671
Db 2068 sArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLy 2088
Qy 6672 AGGACCTGTCATAGGAAGGTTCCCCATCCCAACACCTTTGTAAACACACTGAAGC 6731
Db 2088 sGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCysTyrglySerThrValVa 2108
Qy 6732 CTTTAATCAAACTCTCAGAGTCAATATCTGGTACCTGGTTATGACAGGTGGTGGCATG 6791
Db 2108 lTyrgluGlnAlaSerGluAsnGlnTyrglyLeuValProGlyTyrglnAsnArgTrpAlaCy 2128
Qy 6792 TAATACTGGATTAAACCTCTGTGTTCACCTTGTGTTTTTAAACCAAACTAAAGATTTTTG 6851
Db 2128 sAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCy 2148
Qy 6852 CATTATGGTCAAAATTTGTTCCCGAGTGATTACTATCCGAAAAAGCAATCCTTGTATGA 6911
Db 2148 sValMetValGlnIleValProArgValTyrglyHisProGluGluValValLeuAspG 2168

Qy	6912	ATATGACTACAGAAATCATGACACAAAGAGAGAACCCCATATCTCTGACACTTGTCTGTGAT	6997
Db	2168	uTyAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMe	2188
Qy	6972	GCTCGACTGCGAGTGGCAGCAGGCTGTAGGAACACAGGACAGCTGCCCTGTCTCAGGGACC	7031
Db	2188	tLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyPr	2208
Qy	7032	ACAGCAGCTAGAAAACAGGACTTAGTAACTACATCGAATTGTAAACAGAGAATCTCCAAGC	7091
Db	2208	oGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAl	2228
Qy	7092	CCTAGAAAAATCTGTCTAGTAACTGGGAGGAATCCCTAACCTCTTATCTGAAGTAGPCT	7151
Db	2228	aLeuGluGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValValLe	2248
Qy	7152	ACAGAAATGAGAGGGTTAGATTATTTATTTCTTAAAGAGGAGGATTATGTGTAGCCTT	7211
Db	2248	uGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLe	2268
Qy	7212	GAAGGAGGAATGCTGTTTATTTATGTGGCATCATTTACAGGGGCCATCAGAGACTCCCATGAACAA	7271
Db	2268	uLysGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetSerLy	2288
Qy	7272	GCTTAGAGAAAGTTGGAGAAAGCCTCGAAGGAAAAGAACTACTCAAGGGTGGTTTGA	7331
Db	2288	sLeuArgGluArgLeuGluArgAArgGluArgGluAlaAspGlnGlyTrpPheGln	2308
Qy	7332	GGGATGCTTCAACAGGCTCTCTTTGGTGGCTACCCCTACTTCTTGCTTTAACAGACCCCTT	7391
Db	2308	uGlyTyrPheAsnArgSerProTyrMetThrThrLeuLeuSerAlaLeuThrGlyProLe	2328
Qy	7392	AATAGTCTCTCTCTCTTACTTCACAGTTGGGCCATGTATTATTAAACAAGTTAATTGCTT	7451
Db	2328	uValValLeuLeuLeuLeuLeuThrValGlyProCysLeuLeuAsnArgPheValAlaPh	2348
Qy	7452	CATTAGAGAACGAATAGTCAGTCCAGATCATGTGTACTTAGACAAACAGTACCAC	7506
Db	2348	eValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGln	2366
RESULT 3			
Q9Q1X3_9GAMR			
ID	Q9Q1X3_9GAMR	PRELIMINARY;	PRT; 2376 AA.
AC	Q9Q1X3;		
DT	01-MAY-2000	(TrEMBLrel. 13, Created)	
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Type C proviral gag, pol and env genes and LTR (class B, clone 43).		
OS	Porcine endogenous retrovirus.		
OC	Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;		
OX	Gammaretrovirus; 1-Mammalian type C virus group.		
NCBI_Taxid=61673;			
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Type C;		
RX	MEDLINE=20219394; PubMed=10756014;		
RX	DOI=10.1128/JVI.74.9.4028-4038.2000;		
RA	Czauderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;		
RT	"Establishment and characterization of molecular clones of porcine		
RT	endogenous retroviruses replicating on human cells.,"		
RL	J. Virol. 74:4028-4038(2000).		
DR	EMBL;	AJ133818;	CAB65341.1; -; Genomic_DNA.
DR	HSP;	P03385;	IMOF.
DR	SMR;	Q9Q1X3;	7-58; 2225-2277.
DR	MEROPS;	A02.020;	-.
DR	GO;	GO:0019028;	C: viral capsid; IEA.
DR	GO;	GO:0004190;	F: aspartic-type endopeptidase activity; IEA.
DR	GO;	GO:0003677;	F: DNA binding; IEA.
DR	GO;	GO:0004523;	F: ribonuclease H activity; IEA.
DR	GO;	GO:0003723;	F: RNA binding; IEA.
DR	GO;	GO:0003964;	F: RNA-directed DNA polymerase activity; IEA.
DR	GO;	GO:0005198;	F: structural molecule activity; IEA.

1065 CCCTTTGCCCTCTGGAGCTCCGCCGTGGAGGACCTGCTGCAGGACCTCGAGCCGG 1124
Db ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
1125 AGGGCGCCACCCCGGAGCGACAGACAGAGATCGCAGATTCACCGCTCGCGACAGTACGGC 1184
Db ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
1185 CCTCCACACCGGGGGGCAATGTCAGCCCTCCAGTATGCGCCCTTTCTTCGTCAGAT 1244
Db ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220
1245 CTCTATAATTCGAAAACTAACCATCCCTTCTCGGAGGATCCCAACGCTCACCGGG 1304
Db LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
1305 TTGGTGGAGTCCCTTATGTTCTCTCACCGAGCTACTTGGGATGATTGTCAACAGCTGCTG 1364
Db LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
1365 CAGACACTCTTCACACCGAGAGGGAGAGAAATTTCTATTAGAGGCTAGAAAAATGTT 1424
Db GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaAArgLysAsnVal 280
1425 CCTGGGCGGACGGCGCACCCACCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTA 1484
Db ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
1485 ACTGCCCGGTTGGGACTACAAACACGCTGAAAGTAGGAGAGCTTGAAAATCTATCGC 1544
Db ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
1545 CAGGCTCTGGTGGCGGGTCTCGGGCGCCTCAAGACGCGCCACTAATTTGGCTAAGGTA 1604
Db GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
1605 AGAGAGTGATGCGAGGACCAATGAAACCCCTCTGTTTCTTGAGAGCTCTTGAA 1664
Db ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
1665 GCCTTCAGCGGTACACCCCTTTTGATCCCACTCAGAGGCGCAAAAGCCTCAGTGGCT 1724
Db AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
1725 TTGGCTTTTATGACAGCTCAGCCTTGATATTTAGAAAGAGCTTCAGAGACTGGAAGGG 1784
Db LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
1785 TTACAGGAGGCTGAGTTACGTGATCTAGTGAAGGAGGAGAGAAAGTATATTACAAAAG 1844
Db LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 420
1845 GAGACAGAAGAAAGGAAACAAAGAAAAGAGAGAGAAAGAGAAAGGAGGAAAGA 1904
Db GluThrGluGluGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 440
1905 CGTAATAACCGCGAAGAGAAGAAATTTGACTAAGATCTTGGCTGCGAGTGGTTGAAGGAAA 1964
Db ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLys 460
1965 AGCAATACGGAAGAGAGAGATTTTAGGAAATTAGTGACGCGCCCTAGACAGTCAGGG 2024
Db SerSerArgGluArgGluArgPheArgLysIleArgSerGlyProArgGlnSerGly 480
2025 AACCTGGCAATAGACCCCACTGCACAAGACCAATGTGCATATTTGTAAGAAGAGAGA 2084
Db AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
2085 CACTGGGCAAGGAATGCCCCAAGAGGAAACAAAGACCAAGGATCTTAGTCTAGAA 2144
Db HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520

2145 GAAGATAAAGATTAGGGAGACGGGTTTCGAGCCCTCCCGAGCCCGAGGTAACTTTG 2204
Db GluAspLysAsp---GlyArgArgGlySerAspProLeuProGluProArgValAlaLeu 539
2205 AAGTGGAGGGCAACAGATTGAGTTCTCTGGTGTATACCGGAGCGAAACATTTCAGTGCTA 2264
Db LysValGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeu 559
2265 CTACAGCCATTAGAAAACATAAAGATATAAATACTCTGGGTGATGGTCCACAGGGCAA 2324
Db LeuGlnProLeuGlyLysLeuLysLysSerTrpValMetGlyAlaThrGlyGln 579
2325 CAACAGTATCCATCGACTACCCGAAGAACAGTTGACTTGGGAGTGGGACGGTAAACCCAC 2384
Db ArgGlnTyrProTrpThrThrArgArgThrValAspLeuAlaValGlyArgValThrHis 599
2385 TCGTTTCTGGTCATACCTGAGTCCCGACCCCTCTTTAGGTAGAGACTTATTGACCAAG 2444
Db SerPheLeuValIleProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLys 619
2445 ATCGGAGCACAAATTTCTTTTGAACAAGGGAACCAAGAAGTGTCTGCAAAATAACAAACCT 2504
Db MetGlyAlaGlnIleSerPheGluGlnGlyArgProGluValSerValAsnAsnLysPro 639
2505 ATCACTGTGTGACCCCTCCAAATTAGATGACAAATATCGACTATATCTCTCCCTAGTAAAG 2564
Db IleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnValLys 659
2565 CCTGATCAAAATATACAAATTCGTGTGGAACAGATTTCCCAAGCTCGGGCAGAAACCGCA 2624
Db ProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAla 679
2625 GGGATGGGTTGGCAAGCAGATTTCCCCCACTGATTTTCACTGAGAGCCAGTGCACACA 2684
Db GlyMetGlyLeuAlaLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThr 699
2685 CCAGTGTGAGTCAGACAGTACCCCTTCAGTAAAGAAAGCTCAAGAAAGAAATTCGGCCGAT 2744
Db ProValSerValArgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHis 719
2745 GTCCAAAGATTAAATCCCAACAGGGCATCTAGTTCTCTGTCCAAATCTCCCTGGAATCTCCC 2804
Db ValGlnArgLeuIleGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrPro 739
2805 CTGTACTCGGTTAGAAAAGCCTGGGACTAATGACTATCGACCAGTACAGGACTTCGAGAG 2864
Db LeuLeuProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGlu 759
2865 GTCAATAAACCGGTGCGAGGATATACCCCAACAGTCCCGAACCCCTTATAACCTCTTGTGT 2924
Db ValAsnLysArgValGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSer 779
2925 GCTCTCCACCCCAACCGGAGCTGTATACATATTGACTTAAAGGATGCTTCTTCTGCG 2984
Db AlalaLeuProProGluArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPheCys 799
2985 CTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAATCGAGAGATCCAGGTACG 3044
Db LeuArgLeuHisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThr 819
3045 GGAAGAACCGGCGAGCTCACCTGGACCCGAGCTGCGCCCAAGGGTTCAAGAACTCCCGGACC 3104
Db GlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThr 839
3105 ATCTTTGACGAAGCCCTACACAGAGCTGGCCCAATTCAGGATCCCAACACCTCAGGTG 3164
Db IlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnVal 859
3165 ACCCTCTCCAGTACGTGGATGACTCTCTCTGGGGGAGCCCAACACAGAGCTGCTTA 3224
Db ThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThrLysGlnAspCysLeu 879
3225 GAAGGCACGAAGGCACACTACTGCTGGAATTTGCTGACCTTAGGCTACAGAGCCTCTGCTAAG 3284

Db 880 GluGlyThrLysAlaLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLys 899
QY 3285 AAGCCCAAGATTTCAGGAGAGAGTAACTATTTGGGGTACAGTTTGGCGGACGGCAG 3344
Db 900 LysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGln 919
QY 3345 CGATGGCTGACGGAGGACGGAAGAACTCTAGTCCAGATACCGGCCCCAACACACAGCC 3404
Db 920 ArgTrpLeuThrGluAlaArgLysLysThrValValGlnIleProAlaProThrAla 939
QY 3405 AAACAATAGAGAGTTTTCGGGACAGCTGGATTTGCAGACTGTGGATCCCGGGTTT 3464
Db 940 LysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPhe 959
QY 3465 GCGACCTTAGCAGCCCACTTACCCGCTAACCAAGAAAAAGGGGAATTCCTCGGGCT 3524
Db 960 AlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGlyGluPheSerTrpAla 979
QY 3525 CCTGAGCACCAAGGCAATTTGATGTATCAAAAAGGCCCTGTGAGCGCACCTGTCTCTG 3584
Db 980 ProGluHisGlnLysThrPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeu 999
QY 3585 GCCCTCCTGACGTAAACCCCTTACCTTTATGTGGATGAGCTAAGGAGTAGCC 3644
Db 1000 AlaLeuProAspValThrLysProPheThrLeuTyrValAspGluArgLysGlyValAla 1019
QY 3645 CGGGAGTTTAAACCCCAACCTAGGACCATGGAGAGACCTGCGCTACCTGTCAAG 3704
Db 1020 ArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLys 1039
QY 3705 AAGCTCGATCCTGTAGCAGTGGTGGCCCATATGCTGAAGGCTATCGCAGCTGTGGCC 3764
Db 1040 LysLeuAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaValAla 1059
QY 3765 ATACTGTCAAGACGCTGACAAATTGACTTTGGGACAGATATACTGTAATAGCCCCC 3824
Db 1060 IleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaPro 1079
QY 3825 CATGCTATTGAGAACATCGTTTCGGCAGCCGCCAGACCGATGGATGACCAACGCCGATG 3884
Db 1080 HisAlaLeuGluAsnIleValArgGlnProAspArgTrpMetThrAsnAlaArgMet 1099
QY 3885 ACCCACTATCAAGCCTGTCTTCCAGAGAGGGTCACTGCTCGCTCCACAGCGCTCTC 3944
Db 1100 ThrHisTyrGlnSerLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeu 1119
QY 3945 AACCTCGCACTCTTTCGCTGAAGAGACTGATGAACCACTGACTCATGATTCGCATCAA 4004
Db 1120 AsnProAlaThrLeuLeuProGluGluThrAspGluProValThrHisAspCysHisGln 1139
QY 4005 CTATTGATTGAGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAA 4064
Db 1140 LeuLeuIleGluThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGlu 1159
QY 4065 GTGCTAACCTGGTTTCACTGACGGAAGCAGCTATGTGTGGAAGGTAAAGAGATGGCTGGG 4124
Db 1160 ValLeuThrTrpPheThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGly 1179
QY 4125 GCGCGGTGTGACGGGACCCGACGATCTGGGCGAGCAGCTGCGCGGAAGCACTTCA 4184
Db 1180 AlaAlaValValAspGlyThrHisThrIleTrpAlaSerSerLeuProGluGlyThrSer 1199
QY 4185 GCACAAAGGCTGAGCTCATGGCCCTCACCGACCTTGGCGCTGGCCGAGGGAATCC 4244
Db 1200 AlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSer 1219
QY 4245 ATAAACATTTATACGGACACAGGTATGCTTTGCGACTCCACACGTACATGATGGGCCATC 4304
Db 1220 IleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyValIle 1239
QY 4305 TATAAACAAGGGGTTGCTTACTCAGCAGGAGGGAATAAAGAACAAAGAGGAAAT 4364

Db 1240 TyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIle 1259
QY 4365 CTAAGCCTATTAGAAGCCGTACATTTTACAAAAGGCTAGCTATTATACACTGTCCTGGA 4424
Db 1260 LeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGly 1279
QY 4425 CATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATGGCTGACCGGGTTCGCAAG 4484
Db 1280 HisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLys 1299
QY 4485 CAGGACCCCGAGGTGTTAACTCTTCCTTATATAGAAATGCCCAAGCCCGCAAGACCC 4544
Db 1300 GlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThrProLysAlaProGluPro 1319
QY 4545 AGACGACAGTACACCCCTAGAAAGACTGGCAAGAGATAAAAAGATAGACACCTCTCTGAG 4604
Db 1320 ArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGlu 1339
QY 4605 ACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATCTGCCCCCAACAAGAGGG 4664
Db 1340 ThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGly 1359
QY 4665 TTAGATATATGTCACACAGATACATGCTTAACCCCTAGAACTAAACACCTGCGACAG 4724
Db 1360 LeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGln 1379
QY 4725 TTGCTCAGAACATCCCTTATCATGTTCTCAGAGCTACAGAGTGGCTGACTCGTGTGC 4784
Db 1380 LeuValArgThrSerProTyrHisValLeuArgLeuProGlyValAlaAspSerVal 1399
QY 4785 AAACATTTGTGTCGCCCTGCCAGCTGGTTAATGCTTAATCTTCCAGAAATCCCTCCAGGAAG 4844
Db 1400 LysHisCysValProCysGlnLeuValAsnAlaAsnProSerArgValProProGlyLys 1419
QY 4845 AGACTAAGGGGAGGACACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCG 4904
Db 1420 ArgLeuArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluValLysPro 1439
QY 4905 GCTAAATACGGAACAAATACCTATTGTTTGTAGACACCTTTTTCAGATCGGTAGAG 4964
Db 1440 AlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGlu 1459
QY 4965 GCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTTAAATCTTCCAGAAATCTTCCAGGAAG 5024
Db 1460 AlaTyrProThrLysLysGluThrSerThrValValAlaLysLysIleLeuGluIle 1479
QY 5025 TTTCCAAAGATTGGAAATACCTAAGGTAAATAGGTGACAAATGGTCAGACTTTTGTGGC 5084
Db 1480 PheProArgPheGlyIleProLysValIleGlySerAspAsnGlyProAlaPheValAla 1499
QY 5085 CAGTTAAGTCAGGACTGGCCAAAGATATTGGGATTTGATTGGAAACTGCAATTTGTCATAC 5144
Db 1500 GlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyr 1519
QY 5145 AGACCCCAAGCTCAGGACAGGTAGAGAGATGAATGAACCACTTAAAGAGACCTTTACT 5204
Db 1520 ArgProGlnSerSerGlyGlnValGluArgMetAsnLysThrIleLysGluThrLeuThr 1539
QY 5205 AAATTGACCGCGAGACTGGCGTTAATGATTTGATAGTCTCTCTGCGCTTTGTGCTTTT 5264
Db 1540 LysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeuLeuProPheValLeuPhe 1559
QY 5265 AGGGTTAGGACACCCCTGGACAGTTTGGCTGACCCCTATGAATTTACTCTACGGGGA 5324
Db 1560 ArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGlnLeuLeuTyrGly 1579
QY 5325 CCCCCCACTTGTAGAAATTTGCTTGTACATAGTGTGCTGCTGCTTTTCCAGCCCT 5384
Db 1580 ProProProLeuValGluIleAlaSerValHisSerAlaAspValLeuLeuSerGlnPro 1599
QY 5385 TTGTTCTCTAGGCTCAAGGCACTTGAGTGGGTGAGACACGAGCGTGGAGGCACTCCGG 5444
Db 1600 LeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArg 1619

```
QY 5445 GAGGCTACTCAGGAGGAGAGACTTGCGATGCCACATCGTTTCCAAAGTGGGAGATTCA 5504
Db 1620 GluAlaTyrSerGlyGluGlyAspLeuGlnValProHisArgPheGlnValGlyAspSer 1639
QY 5505 GTCTACGTTAGACGCCACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGGCCCTTATCTC 5564
Db 1640 ValTyrValArgHisArgAlaGlyAsnLeuGluThrArgTTrpLysGlyProTyrLeu 1659
QY 5565 GTACTTTTGACCACACCGCTGTGAAAGTCGAAGATCTCCACTGATCCATGCCA 5624
Db 1660 ValLeuLeuThrThrProThrAlaValLysValGluGlyIleSerThrTrpIleHisAla 1679
QY 5625 TCCACAGCTTAAACCGCGCCACCTCCGATTTCGGGTGCGAAGCCGAAAGACACTGAAAT 5684
Db 1680 SerHisValLysLeuAlaProProAspSerGlyTrpArgAlaGluLysThrGluAsn 1699
QY 5685 CCCCTTAAGCTTCGCCCTCCATCGGTGTTCTTACTCTGTCAATAACCTCT----- 5736
Db 1700 ProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsnAsnAsn-SerProGlyG1 1719
QY 5736 ----- 5736
Db 1719 nMethIleProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluProLysArgLe 1739
QY 5737 -----CA 5738
Db 1739 uArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleThrProG1 1759
QY 5739 GACTAATGGTATGGCATPAGGAGACAGCCTGAACTCCCATAAACCTTATCTCTCACTG 5798
Db 1759 nAlaSerSerLysArgLeuIleAspSerSerAsnProHisArgProLeuSerLeuThrTr 1779
QY 5799 GTTAATTACTGATCCGGCAAGGTATTAATATCAACAACTCAAGGGAGGCTCCTTT 5858
Db 1779 pLeuIleIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProAr 1799
QY 5859 AGGAACCTGGGCTGATCTATACGTTTGCTCAGACTCAGTTATCTCTAGCTCTGACTC 5918
Db 1799 gGlyThrTrpTrpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSe 1819
QY 5919 A---CCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCACCAATAA 5975
Db 1819 rThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyThrGluLysG1 1839
QY 5976 TGGAAACATTCGGGAATCCAGAGATTTCTTTTGTAAACAACTGGAACCTGTAACCTC 6035
Db 1839 u---LysTyrCysGlyGlySerGluGluSerPheCysArgArgTrpSerCysValThrSe 1858
QY 6036 TAATGATGATATTGGAAATGCCAACCTCTCAGCAGGATAGGTTAAGTTTCTTATGT 6095
Db 1858 rAsnAspGlyAspTrpLysTrpPheIleSerLeuGlnAspArgValLysPheSerPheVa 1878
QY 6096 CAACACCTATACAGCTCTGGACAATTTAAATTACTCTGACCTGGATTAGAACTGGAAGCCC 6155
Db 1878 lAsn-----SerGlyProGlyLysTyrLysValMetLysLeuTyrLysAspLy 1894
QY 6156 CAAGTGCTCTCTCAGACCTAGATTACTTAAATAAGTTTCACTGAGAAGGAAACA 6215
Db 1894 sSerCysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysG1 1914
QY 6216 AGAAATATCTAAATGGTAAATGGTATCTCTCTGGGAATGGTATAT-----TATGG 6269
Db 1914 nGluAsnIleGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrG1 1934
QY 6270 AGGCTCGGGTAAACAACACGAGCTCCATTCTCACTATTCCGCTCAAAATAAACACCAG---CT 6326
Db 1934 yGlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyTh 1951
QY 6327 GGAGCTCTCAATGGCTATAGGACCAATACGGTCTTGACGGGTCAAGACCCCCCAACCCA 6386
Db 1951 rGluProProValAlaValGlyProAspLysValLeuAlaGluGlnGlyProProAlaLe 1971
```

```
QY 6387 AGGACCA-----GG 6395
Db 1971 uGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrG1 1991
QY 6396 ACATCTCTTAACATACTTCTGGATCAGACCCACACTGATGCT-----AACAGC----- 6444
Db 1991 nProProSerAsnGlyThrThrGlyLeuIleProThrAsnThrProArgAsnSerProG1 2011
QY 6445 ----ACGACTAAATGGGGCAAAACTTTTAGCTCATCCAGGAGCTTTTCAAGCTCT 6500
Db 2011 yValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaI1 2031
QY 6501 TAACTCCACGACTCCAGAGGCTACTCTTCTTGTGGCTATGCTTAGCTTCGGGCCACACC 6560
Db 2031 eAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPr 2051
QY 6561 TTACTATGAAGGAATGCTAGAGAGGAAATCAATGTGACAAAAGAACATAGAGACCA 6620
Db 2051 oTyrTyrGluGlyMetAlaLysGluGlyLysPheAsnValThrLysGluHisArgAsnG1 2071
QY 6621 ATGCACATGGGATCCCAAAATTAAGCTTACCTTACTAGGTTTCTCGAAAAGGCACCTG 6680
Db 2071 nCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCy 2091
QY 6681 CATAGGAAGGTTCCCCATCCCAACCAACACTTTGTAACCACTGAAGCCTTTAATCA 6740
Db 2091 sIleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluG1 2111
QY 6741 AACCTCTGAGAGTCAATATCTGTPACTGTTATGACAGGTGGTGGCATGTAATACTGG 6800
Db 2111 nAlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrG1 2131
QY 6801 ATTAACCCCTGTCTTCCACCTTGGTTTAAACCAAACTAAAGATTTTTCATATTAGGT 6860
Db 2131 yLeuThrProCysValSerSerValPheAsnGlnSerLysAspPheCysValMetVa 2151
QY 6861 CCAATTTGTTCCCGAGTGTATTACTATCCCGAAAAAGCAATCCTTGATCAATATGACTA 6920
Db 2151 lGlnIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAspTy 2171
QY 6921 CAGAAATCATCGACAAAGAGAGAACCCATATCTCTGACACTTGTCTGTGATGCTCGGACT 6980
Db 2171 rArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLe 2191
QY 6981 TGGAGTCGACGAGCTCTAGCAACAGCAACAGCTGCCCTCGTACCGGACCCACAGCAGCT 7040
Db 2191 uGlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLe 2211
QY 7041 AGAAACAGGACTTAGTAACCTTACATCCAATTTGTAACAGAAGATCTCCAAGCCCTAGAAAA 7100
Db 2211 uGluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuGluG1 2231
QY 7101 ATCTGTCAAGTAACCTGGAGGAATCCCTTAACCTCTTATCTGAAGTAGTCTCTACAGATAG 7160
Db 2231 uSerValSerAsnLeuGluSerLeuThrSerLeuSerGluValValLeuGlnAsnAr 2251
QY 7161 AAGAGGTTAGATTATTTATTTCTAAAGAGGAGGATTATGCTGTAGCCTTGAAGGAGGA 7220
Db 2251 gArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluG1 2271
QY 7221 ATGCTGTTTTTATGTGATCATTCAGGGGCCATCAGAGACTCCCATGAACAGCTTAGAGA 7280
Db 2271 uCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetSerLysLeuArG1 2291
QY 7281 AAGGTTGAGAACGCTCGAAGGGGAAAAAGGAAACTACTCAAGGGTGGTTTGAAGGATGGTT 7340
Db 2291 uArgLeuGluArgArgArgGluArgGluAlaAspGlnGlyTrpPheLysGlyTrpPh 2311
QY 7341 CMACAGCTCTCTTGGTTGGCTACCTTCTTCTTCTTAAACAGGACCTTAAATAGTCCT 7400
Db 2311 eAsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLe 2331
QY 7401 CCTCTGTACTCACAGTTGGGCCCATGTATTATTAAACAAGTTAATTCCTTTCATTAGAGA 7460
```


Db 2331 uLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGI 2351
Qy 7461 ACGAATAAGTCAGTCAGATCATGGTACTTAGACAAACAGTACCAG 7506
Db 2351 uArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGln 2366

RESULT 4

ID Q8J4V6_9GAMR PRELIMINARY; PRT; 1720 AA.
AC Q8J4V6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
OS Gag-pol polyprotein.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=194959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22172947; PubMed=12185278;
RA Bartosch B., Weiss R.A., Takeuchi Y.;
RT "PCR-based cloning and immunocytological titration of infectious
RT porcine endogenous retrovirus subgroup A and B.";
RL J. Gen. Virol. 83:2231-2240(2002).
DR EMBL; AY099324; AAM29194.1; -; Genomic_DNA.
DR HSSP; P03355; 116J.
DR SMR; Q8J4V6; 7-98.
DR MEROPS; A02_020; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00879; RNASE_H; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.
KW Polyprotein.
SQ SEQUENCE 1720 AA; 193393 MW; 7F9DCC2F5E834D8A CRC64;

Alignment Scores:

Pred. No.: 0 Length: 1720
Score: 8821.00 Matches: 1655
Percent Similarity: 98.4% Conservative: 34
Best Local Similarity: 96.4% Mismatches: 27
Query Match: 60.3% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3 (1-8132) x Q8J4V6_9GAMR (1-1720)

Qy 585 ATGGGACAGACGGTGCAGACCCCTCTTGTAGTTTGACTCTCGACCATTCGAGTTAAAG 644
Db 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTTrpThrGluValArg 20
Qy 645 TCAGGGCTCATTAATTGTCTCAGTTACAGTTAAGAGGGACCTTGGCAGACTTCTGTGTC 704
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysGlyProTTrpGlnThrPheCysAla 40
Qy 705 TCTGAATGGCCGACATTCGATGTGGATGCCCATCAGAGGGAGCCTTTAATCTGAGATT 764
Db 41 SerGluTTrpProThrPheAspValGlyTTrpProSerGluGlyThrPheAsnSerGluIle 60
Qy 765 ATCTGTGCTTTAAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCGATCAGAG 824
Db 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
Qy 825 CCCTATATCTTACGTGGCAAGATTTGGCAGAGAGATCTCCGCCATCGGTTAAACATGG 884
Db 81 ProTyrIleLeuThrTTrpGlnAspLeuAlaGluAspProProTTrpValLysProTTrp 100
Qy 885 CTGAATAAGCCAAGAACCCAGTCCCGCAATCTTGGCTCTTCGAGAGAGAAAACAAACAC 944
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
Qy 945 TCGGCTGAATAAGTCAAGCCCTCTCTCATATCTTACCCCGAGATTGAGGAGCCACCGGCT 1004
Db 121 SerAlaGluLysValGluProSerProArgIleTyrProGluIleGluGluProThr 140
Qy 1005 TGGCCGGAACCCCAATCTGTTCCTCCACCCCTTATTCGCACAGGTGGCGGAGGGA 1064
Db 141 TrpProGluProGlnProValProProProTTrpProAlaGlnGlyAlaValArgGly 160
Qy 1065 CCCTTTGCCCTCTCTGGAGCTCCGGGCTGGAGGAGCTCTCGCAGGACTCGAGGCGG 1124
Db 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
Qy 1125 AGGGGCGCCACCCCGGAGCGGACAGAGATCGCAGATTACCGCTGCGCACGTACGGC 1184
Db 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
Qy 1185 CTCTCCACACCGGGGGGCGCAATTGCGACCCCTCCAGTATTTGGCCCTTTCTCTGAGAT 1244
Db 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTTrpProPheSerSerAlaAsp 220
Qy 1245 CTCTATAATTGGAAACTAACCATCCCTTTCTCGGAGGATCCCAACCCCTCACCGGG 1304
Db 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
Qy 1305 TTGGTGGAGTCCCTTATGTTCTCTCACAGCCTTACTTGGGATGATTGTCAACAGCTCTGT 1364
Db 241 LeuValGluSerLeuMetPheSerHisGlnProThrTTrpAspAspCysGlnLeuLeu 260
Qy 1365 CAGACACTTTCACAAACCGGAGCGGAGAGAGAAATTTCTATTAGAGGCTAGAAAAATGTT 1424
Db 261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
Qy 1425 CCTGGGGCGGACGGCGACCCCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTA 1484
Db 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
Qy 1485 ACTCGCCCGGTTGGGACTACAAACAGCGCTGAAAGGTAGGAGAGCTTGAATAATCATCGC 1544
Db 301 ThrArgProGlyTTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
Qy 1545 CAGGCTCTGGTGGCGGGTCTCCGGGGCGCTCAAGACGGCCCACTAATTTGGTGAAGGTA 1604
Db 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
Qy 1605 AGAGNAGTGTATCGAGGACCGGATGAACCCCTCTGTTTCTTCTGAGAGGCTCTTGAA 1664
Db 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360

QY 1665 GCCTTCAGGGGTACACCCCTTTTGATCCACCTCAGAGGCCCAAAAGCCCTCAGTGGCT 1724
Db |||||:::|||||
361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
QY 1725 TTGGCCCTTTATAGACACAGTCCAGCCCTTGATATTAGAAAAGAGCTTCAGAGACTCGAAGGG 1784
Db |||||:::|||||
381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
QY 1785 TTACAGAGGCTGAGTTACGTGATCTAGTGAAGGAGGACGAGAAAGTATATTACAAAGG 1844
Db |||||:::|||||
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 420
QY 1845 GAGACAGAAAGAAAGGAAACAAGAAAGAGAGAGAAAGAGAGGAAAGGAGAAAGA 1904
Db |||||:::|||||
421 GluThrGluGluGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 440
QY 1905 CGTAATAAACCGCAAGAGAGAAGAAATTGACTAAGATCTTGCTGCAGTGGTTGAAAGGAAA 1964
Db |||||:::|||||
441 ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLys 460
QY 1965 AGCAATACGGAAGACAGAGAGAGATTTTAGGAAATTAGCTCAGGCCCTAGACAGTCAAGG 2024
Db |||||:::|||||
461 SerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
QY 2025 AACCTGGGCAATAGGACCCCACTCGACAAGACCAGTGTGCATATTGTAAAGAAAGAGGA 2084
Db |||||:::|||||
481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
QY 2085 CACTGGGCAAGAACTGCCCAAGAGGAAACAAGGACCAAGAGCTCTAGCTCTAGAA 2144
Db |||||:::|||||
501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
QY 2145 GAAGATAAAGATTAGGGGAGACGGGTTTCGAGCCCTCCCGAGCCGAGGGTAACTTTG 2204
Db |||||:::|||||
521 GluAspLysAsp**GlyArgArgGlySerAspProLeuProGluProArgValThrLeu 540
QY 2205 AAGGTGGAGGGCAACAGATTGAGTCTCGTCTGCATACCGGAGCGAAACATTCAGTGCTA 2264
Db |||||:::|||||
541 LysValGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeu 560
QY 2265 CTACAGGCATTAGGAAAATAAAAGATAAAATCTCGGTGATGGGTGCCACAGGGCAA 2324
Db |||||:::|||||
561 LeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMetGlyAlaThrGlyGln 580
QY 2325 CAACAGTATCCATCGACTACCCGAAGACAGTTGAGCTTGGAGTGGGACGGGTAAACCCAC 2384
Db |||||:::|||||
581 ArgGlnTyrProTrpThrThrArgArgThrValAspLeuGlyValGlyArgValThrHis 600
QY 2385 TCGTTTCTGGTTCATACCTGAGTGCACAGCCCTCTTAGGTAGAGACTTATTCCACCAAG 2444
Db |||||:::|||||
601 SerPheLeuValIleProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLys 620
QY 2445 ATGGGAGCACAAATTTCTTTTGAACAAGGAAACGAGAAGTGTCTGCAAAATAACAAACCT 2504
Db |||||:::|||||
621 MetGlyAlaGlnIleSerPheGluGlnGlyArgProGluValSerValAsnAsnLysPro 640
QY 2505 ATCACTGTGTGACCTCCCAATTAGATGACGAATATCGACTATACTCTCCCTAGTAAAG 2564
Db |||||:::|||||
641 IleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnValLys 660
QY 2565 CCTCATCAAAATATACAAATTCGTGTTGGAACAGTTTCCCAAGCTCGGGCAGAAACCGCA 2624
Db |||||:::|||||
661 ProAspGlnAspIleGlnSerTrpPheGluGlnPheProGlnAlaTrpAlaGluThrAla 680
QY 2625 GGGATGGGTTTGGCAAGCAAGTTTCCCCCAACAAGTTATTCAACTGAAGGGCCAGTGCCACA 2684
Db |||||:::|||||
681 GlyMetGlyLeuAlaLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThr 700
QY 2685 CCAGTGTCAAGTACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAAATTCGGCCCGCAT 2744
Db |||||:::|||||
701 ProValSerValArgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHis 720
QY 2745 GTCCAAAGATTAACTCAACAGGGCATCCTAGTCTCTGTGCCAATCTCCCTGGAACTACTCCC 2804

Db |||||:::|||||
721 ValGlnArgLeuIleGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrPro 740
QY 2805 CTGCTACCGGTTAGAAAGCCTGGGACTTAATGACTATTCGACACAGTACAGGACTTGAGAGAG 2864
Db |||||:::|||||
741 LeuLeuProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGlu 760
QY 2865 GTCAATAAAGCGGTGCGAGGATATACACCCAAACAGTCCCGAACCTTATAACCTCTCTGT 2924
Db |||||:::|||||
761 ValAsnLysArgValGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSer 780
QY 2925 GCTCTCCACCCCAACGAGCTGCTATACAGTATTGACTTAAAGAGTGCCTCTCTTCG 2984
Db |||||:::|||||
781 AlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCys 800
QY 2985 CTGAGATTACACCCCACTAGCCAAACCACTTTTGGCTTCGAATGGAGAGATCCAGGTACG 3044
Db |||||:::|||||
801 LeuArgLeuHisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThr 820
QY 3045 GGAAGAACCGGGCAGCTCACCTGGACCCGACTGCCCAAGGTTTCAAGAACTCCCCGACC 3104
Db |||||:::|||||
821 GlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThr 840
QY 3105 ATCTTTGACGAAGCCCTACACAGAGACTCTGGCCAACTTCAGGATCCAAACCCCTCAGGTG 3164
Db |||||:::|||||
841 IlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnVal 860
QY 3165 ACCCTCTCCAGTACGTGGATGACCTGCTCTGCGGGAGCCACCACCAAGAGGAGTGTCTTA 3224
Db |||||:::|||||
861 ThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAlaThrLysGlnAspCysLeu 880
QY 3225 GAAGGCACGAAGCAGCTACTGCTGGATTCCTGACCTAGGCTACAGAGCTCTGCTAAG 3284
Db |||||:::|||||
881 GluGlyThrLysAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 900
QY 3285 AAGCCCCAGATTTGACGAGAGAGGTAAACATCTTGGGTGACAGTTTGCAGGACCGGGCAG 3344
Db |||||:::|||||
901 LysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGln 920
QY 3345 CGATGGCTGACGAGGACCGGAGAAACCTGTAGTCAGATACCGGGCCCAACACAGCC 3404
Db |||||:::|||||
921 ArgTrpLeuThrGluAlaArgLysLysThrValValGlnIleProAlaProThrAla 940
QY 3405 AAACAAATGAGAGAGTTTGGGGACAGCTGGATTTGACAGCTGTGAGTCCCGGGTTT 3464
Db |||||:::|||||
941 LysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPhe 960
QY 3465 GCGACCTTAGCAGCCCCACTCTACCCCTAAACCAAAAGAAAGGGGAATTCCTCTGGGCT 3524
Db |||||:::|||||
961 AlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGlyLysPheSerTrpAla 980
QY 3525 CTTGAGCACGAGAGGATTTGATGCTATCAAAAGCCCTGCTGAGGCGACCTGCTCTG 3584
Db |||||:::|||||
981 ProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeu 1000
QY 3585 GCCCTCCCTCAGCTAAACCCCTTTACCTTTATCTGATGAGCTTAAGGGAGTAGGCC 3644
Db |||||:::|||||
1001 AlaLeuProAspValThrLysProPheThrLeuTyrValAspGluArgLysGlyValAla 1020
QY 3645 CGGGAGCTTTTAAACCCAAACCTTAGGACCATGGAGAAAGACTGTGCGCTACCTCTCAAAG 3704
Db |||||:::|||||
1021 ArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLys 1040
QY 3705 AAGCTCATCTGTAGCCAGTGGTTGGCCCATATGCTTGAAGGCTATCGCAGCTGTGGCC 3764
Db |||||:::|||||
1041 LysLeuAspProValAlaSerGlyTrpProValCysLeuLysAlaIleAlaValAla 1060
QY 3765 ATACTGCTCAAGGACGCTGACAAATTCACCTTTGGGACAGAAATATAACTGTAATAGCCCC 3824
Db |||||:::|||||
1061 IleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaPro 1080
QY 3825 CATGCATTTGGAGAACATCTGTTGGCAGCCCCCAGACCGATGGATGACCAACGCCCGCATG 3884
Db |||||:::|||||

Db 1081 HisAlaLeuGluAanIleValArgGlnProPheArgTrpMetThrAsnAlaArgMet 1100
QY 3885 ACCCACTATCAAGACCTGCTTCTCACAGAGAGGCTACGTTTCGCTCCACAGACCGCTCTC 3944
Db 1101 ThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeu 1120
QY 3945 AACCTGCCACTCTCTCCCTGAAGACAGCTGATCAACAGTGCATCATGATTGCCATCAA 4004
Db 1121 AsnProAlaThrLeuLeuLeuProGluGluThrAspGluProValThrHisAspCysHisGln 1140
QY 4005 CTATTGATTGAGAGACTGGGCTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAA 4064
Db 1141 LeuLeuIleGluThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGlu 1160
QY 4065 GTGCTAACCTGGTTCACTACGGAAGCAGCTATGCTGGAGGTAAAGAGATGGCTGGG 4124
Db 1161 ValLeuThrTrpPheThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGly 1180
QY 4125 GCGCGGTGGTGGAGCGGACCCGACCATCTGGCCAGCTTTGGCTGGCGCGAAGGAAATCC 4244
Db 1181 AlaAlaValValAspGlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSer 1200
QY 4185 GCACAAAGGCTGAGCTCATGGCCCTCACGCAAGCTTTGGCTGGCGCGAAGGAAATCC 4244
Db 1201 AlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSer 1220
QY 4245 ATAAACATTATACGGAAGCAGGTATGCTTTGCGACTGCACACGTACATGGCGGCATC 4304
Db 1221 IleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIle 1240
QY 4305 TATAAACAAAGGGGTGCTTACTCTGACGAGGGAGGAATAAAGAACAAAGAGGAAATT 4364
Db 1241 TyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIle 1260
QY 4365 CTAAGCCTATTAGAAGCGGTACATTTTACCAAAAGGCTAGCTATTATACACTGCTCTGGA 4424
Db 1261 LeuSerLeuLeuGluAlaLeuLeuProLysArgLeuAlaIleIleHisCysProGly 1280
QY 4425 CATCAGAAAGCTTAAAGATCTCATATCCAGAGGAAACAGATGGCTGACCGGGTTGCCAAG 4484
Db 1281 HisGlnLysAlaLysAspLeuIleSerArgLysAsnGlnMetAlaAspArgValAlaLys 1300
QY 4485 CAGGACCCAGGGTGTAACTTCTGCTTATATAGAAATGCCAAGCCCGCCAGAACCC 4544
Db 1301 GlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThrProLysAlaProGluPro 1320
QY 4545 AGACGACGTACACCTTAGAAGCTGCGAAGAGATAAAGATAGACACGTTCTCTGAG 4604
Db 1321 ArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGlu 1340
QY 4605 ACTCCGGAAGGACCTGTATACCTCAGATGGGAAGGAAATCTGCCCCACAAAGAGGG 4664
Db 1341 ThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisGlyGly 1360
QY 4665 TTAGAATATGTCACAGATACATGCTTAACCCACTAGGAACCTAAGACCTGCGAGCAG 4724
Db 1361 LeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGln 1380
QY 4725 TTGCTCAGAACATCCCTTATCATGTTCTGAGGCTACCGAGCTGAGTGCCTGCTGTC 4784
Db 1381 LeuValArgThrSerProTyrHisValLeuArgLeuProGlyValAlaAspSerValVal 1400
QY 4785 AAAATTGTTGTCCTGCCAGCTGGTTAATGCTTAATCTTCCAGAAATGCTCCAGGGAAG 4844
Db 1401 LysHisCysValProCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLys 1420
QY 4845 AGACTAAGGGAAGCCACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCG 4904
Db 1421 ArgLeuArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluValLysPro 1440
QY 4905 GCTAAATACGGAACAAATACCTATTGGTTTTGTAGACACCTTTTCAGATGGGTAGAG 4964
Db 1441 AlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGlu 1460

QY 4965 GCTTATCTACTAAGAAAGAGACTTCAACCGTGGTCTAAAAAATACTGGAAGAAATT 5024
Db 1461 AlaTyrProThrLysLysGluThrSerThrValValAlaLysLysIleLeuGluGluIle 1480
QY 5025 TTTCCAAGATTGGAAATACCTAAGGTAAATAGGGTCAGACAATGTCAGCTTTGTTGCC 5084
Db 1481 PheProArgPheGlyIleProLysValIleGlySerAspAsnGlyProAlaPheValAla 1500
QY 5085 CAGTTAAGTCAGGACTGGCAGAGATATTGGGATTGATTGGAAATGTCATTGTGTCATAC 5144
Db 1501 GlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyr 1520
QY 5145 AGACCCCAAGCTCAGGACAGGTAGAGAGATGAATAAGAACCAATTAAGAGACCTTACT 5204
Db 1521 ArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThr 1540
QY 5205 AAATTGACCGCGAGACTGGCGTTAATGATTGATAGTCTCTCTGCCCTTTGCTGCTTTT 5264
Db 1541 LysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeuLeuProPheValLeuPhe 1560
QY 5265 AGGTTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAAATTTACTCTACGGGGA 5324
Db 1561 ArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrGlyGly 1580
QY 5325 CCCCCCATTGGTAGAAATTGCTTCTGTACATAGTGTGCTGACGTGCTTTTCCAGCCT 5384
Db 1581 ProProProLeuValGluIleAlaSerValHisSerAlaAspValLeuLeuSerGlnPro 1600
QY 5385 TTGTTCTTAGGCTCAAGGCACTGTAGTGGTGAGACCAACAGAGCTGGAGGCACTCCGG 5444
Db 1601 LeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArg 1620
QY 5445 GAGGCTACTCAGGAGGAGAGACTTGCAGATCCCATCGTTTCCCAAGTGGGAGATTCA 5504
Db 1621 GluAlaTyrSerGlyGluGlyAspLeuGlnValProHisArgPheGlnValGlyAspSer 1640
QY 5505 GTCTAGCTTAGACCCACCGTCAGGAAACCTCGAGACTCGGTGGAAGGCGCTTATCTC 5564
Db 1641 ValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeu 1660
QY 5565 GTACTTTTGACCAACCAACCGCTGTGAAGTGAAGAAATCTCCACTGCATCCATGCA 5624
Db 1661 ValLeuLeuThrThrProThrAlaValLysValGluGlyIleSerThrTrpIleHisAla 1680
QY 5625 TCCACAGTTAAACCGCGCCACTCCGATTCGGGTGGAAAGCCGAAAGACTGAAAT 5684
Db 1681 SerHisValLysLeuAlaProProAspSerGlyTrpArgAlaGluLysThrGluAsn 1700
QY 5685 CCCTTTAAGCTTCGCTCCATCGCTGCTTCTTACTCTGCTCAATAAC 5732
Db 1701 ProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsnAsnAsn 1716

RESULT 5

Q8J4V8_9GAMR

ID Q8J4V8_9GAMR PRELIMINARY; PRT; 1718 AA.

AC Q8J4V8;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Gag-pol polyprotein.

OS Porcine endogenous retrovirus A.

OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;

OC 1-Mammalian type C virus group.

OX NCBI_TaxID=194958;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22172947; PubMed=12185278;

RA Bartosch B., Weiss R.A., Takeuchi Y.;

RT "PCR-based cloning and immunocytological titration of infectious

RL J. Gen. Virol. 83:2231-2240(2002).

DR EMBL; AY099323; AAM29192.1; -; Genomic_DNA.

DR HSSP; P03355; 116J.
 DR SMB; O8J4V8; 7-98.
 DR MEROPS; A02.020; -.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR GO; GO:0019068; P:viral assembly; IEA.
 DR InterPro; IPR003036; Gag_P30.
 DR InterPro; IPR000840; G_retro.matrix.
 DR InterPro; IPR001995; Peptidase A2 cat.
 DR InterPro; IPR001969; Pept_Asp_AS_
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVtse.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF01140; Gag_MA; 1.
 DR Pfam; PF02093; Gag_P30; 1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; Rve; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT_1; 1.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS50879; RNASE_H; 1.
 DR PROSITE; PS0158; ZF_CCHC; 1.
 DR Polyprotein.
 KW POLYPEPTIDE.
 SQ SEQUENCE 1718 AA; 193115 MW; 4E86215286C59246 CRC64;

Alignment Scores:
 Pred. No.: 0 Length: 1718
 Score: 8815.50 Matches: 1655
 Percent Similarity: 98.3% Conservative: 33
 Best Local Similarity: 96.3% Mismatches: 29
 Query Match: 60.2% Indels: 1
 DB: 2 Gaps: 1

US-10-723-552-3 (1-8132) x Q8J4V8_9GMR (1-1718)

QY 585 ATGGAGCAGACGGTGGACGCCCTCTTGTGACTCTCGACCATTCGACCTGAAGTTAA 644
 DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20
 QY 645 TCCAGGGCTCATAATTTGTCAAGTTTCAGTTTAAGAGGACCTTCGCGACATTTCTGTGTC 704
 DB 21 SerArgAlaHisAsnLeuSerValGlnValLysGlyProTrpGlnThrPheCysAla 40
 QY 705 TCTGAATGCCGACATTTCGATGGATGGATGGCCATCAGGGGACCTTAATTCGAGATT 764
 DB 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
 QY 765 ATCTGGCTGTAAAGCAGATTATTTTCAGACTCGACCCGGCTCTCATCCCGATCAGGAG 824
 DB 61 IleuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
 QY 825 CCCTATATCCCTTACGTGGCAAGATTTGGCAGAGGATCTCCGCCATGGTTTAAACATGG 884
 DB 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
 QY 885 CTGAATAAGCCAGAACCCAGGTCCTCCGATTTCTGGCTTCCTGGAGGAAAAACAACAC 944
 DB 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
 QY 945 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGACCCCGCT 1004
 DB 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500

DB 121 SerAlaGluLysValGluProSerProArgIleTyrProGluIleGluGluProProThr 140
 QY 1005 TGGCGGGAACCCCAATCTGTTCCCCACCCCTATCTGCGACAGGGTGGCGGAGGGA 1064
 DB 141 TrpProGluProGlnProValProProProProProProProProProProProPro 160
 QY 1065 CCCTTTGCTCCCTCTCGGAGCTCGGCGGTGGAGGACCTGCTGCAGGGAATCGGAGCGG 1124
 DB 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
 QY 1125 AGGGGCGCCACCCCGGAGCGGACAGACAGATCCGACATTCACGTCGCGCACATCGGC 1184
 DB 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
 QY 1185 CTCCCCACACGGGGGCGCCAAATTCGACGCCCTCCAGATTGGCCCTTTCTTCGCGAGAT 1244
 DB 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220
 QY 1245 CTCTATAATTGGAAACCTAACCATCCCTTCTCGGAGGATCCCAACGCTCACCGGG 1304
 DB 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
 QY 1305 TTGGTGGAGTCCCTTATGTTCTCTCACGACCTACTTGGGATGATTCACACAGCTGCTG 1364
 DB 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
 QY 1365 CAGACACTCTTCACACCGGAGGAGCGAGAGAAATTCATATAGAGGCTAGAAAAATGTT 1424
 DB 261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
 QY 1425 CTGGGGCGCAGCGGCGACCCAGCGGTTCGAAATAGAGATTGACATGGGATTTCCCTTA 1484
 DB 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluLeuAspMetGlyPheProLeu 300
 QY 1485 ACTCGCCCGGTTGGGACTACAACACGCTGAAGTTAGGAGAGCTTGAATAATCTATCGC 1544
 DB 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
 QY 1545 CAGGCTCTGTGGGGGTCTCCGGGGCGCTCAAGAGCGGCCACTAATTTGGCTAAGGTA 1604
 DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
 QY 1605 AGAGAGTGTATGCAAGGACCGAATGAACCCCTCTCTGTTTCTTGAGAGCTCTTGAA 1664
 DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
 QY 1665 GCCTTCAGGCGGTACACCCCTTTTGATCCCACTCAGAGGCCCAAAAGCCTCAGTGGCT 1724
 DB 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
 QY 1725 TTGGCCTTTATAGACAGTCAAGCTTCGATATTAGAAGAAGCTTCAGAGCTGGAAGGG 1784
 DB 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
 QY 1785 TTACAGAGGCTGAGTTACGTGATCTAGTGAAGGAGCAGAGAAAGTATATTACAAAGG 1844
 DB 401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 420
 QY 1845 GACAAGAAGAAGGGAACAAAGAAAGAGAGAGAAAGAGAGAAAGAGAGAGAGAGAGAG 1904
 DB 421 GluThrGluGluGluLysGluGlnArgLysGluLysGluLysGluGluArgGluGluArg 440
 QY 1905 CGTAATAAAGCGGAGAGAGAAATTTGACTAAGATCTTGGCTGAGTGGTTGAGGGGAA 1964
 DB 441 ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaValValGluGlyLys 460
 QY 1965 AGCAATACGAAAGAGAGAGAGATTTTAGGAAATATTAGGTTCAGGCCCTAGACATCAGG 2024
 DB 461 SerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
 QY 2025 AACCTGGGCAATAGGACCCCACTCGAAGACCAATGTGCAATTTGTAAAGAAGAGGA 2084
 DB 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500

QY 2085 CACTGGGCAAGGAATCTGCCCAAGAGGGGAAACAAAGGACCAAGGATCCTAGCTTAGAA 2144
DB HistpAlaArgAsnCyseProLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
QY 2145 GAAGATAAAGATTAGGGGAGAGGGGTTCGGACCCCTCCCGAGCCAGGGTAACTTTG 2204
DB GluAspLysAsp**GlyArgGlySerAspProLeuProGluProArgValThrLeu 540
QY 2205 AAGGTGAGGGGCAACCAAGTTGAGTTCTCGTGTGATACCGGAGCGAAACATTTCAGTGCTA 2264
DB LysValGluGlyGlnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeu 560
QY 2265 CTACAGCATTAAGGAAACATAAAGATAAAAAATCTCTGGGTGATGGGTGCCACAGGGCAA 2324
DB LeuGlnProLeuGlyLysLeuLysGluLysLysSerTrpValMetGlyAlaThrGlyGln 580
QY 2325 CAACAGTATCCATGGATACCCGAAGAACAGTTGACCTGGAGTGGGAGCGGTAAACCCAC 2384
DB ArgGlnTyrProTrpThrArgArgThrValAspLeuGlyValGlyArgValThrHis 600
QY 2385 TCGTTTCTGGTCACTACCTGAGTCCAGCACCCCTCTTAGTACAGACTTATTGACCAAG 2444
DB SerPheLeuValIleProGluCyseProValProLeuLeuGlyArgAspLeuLeuThrLys 620
QY 2445 ATGGGAGCACAAATTTCTTTTGAAACAAGGGGAAACAGAAAGTGTCTGCAAAATAACAAACCT 2504
DB MetGlyAlaGlnIleSerPheGluGlnGlyArgProGluValSerValAsnAsnLysPro 640
QY 2505 ATCACTGTGTGACCCCTCCATATAGATGACGAATATGCACTATCTCTCCCTAGTAAAG 2564
DB IleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnValLys 660
QY 2565 CCTGATCAAAATATACAATCTGGTTCGAAACAGTTTCCCAAGCTCGGCGAGAAACCGCA 2624
DB ProAspGlnAspIleGlnSerTrpPheGluGlnPheProGlnAlaTrpAlaGluThrAla 680
QY 2625 GGGATGGTTGGCAAGCAAGTTCCCCCAAGTATTCAACTGAAGGCGAGTGCACCA 2684
DB GlyMetGlyLeuAlaLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThr 700
QY 2685 CCAGTGTCACTCAGACAGTACCCCTTCAGTAAGAAGCTCAAGAGAAATTCGCGCCGAT 2744
DB ProValSerValArgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHis 720
QY 2745 GTCCAAAGATTAAATCCAAAGGGCATCTAGTTCTCTGCTCCAATCTCCCTGGAATACTCC 2804
DB ValGlnArgLeuIleGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrPro 740
QY 2805 CTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACAGTACAGGACTTGAGAGAG 2864
DB LeuLeuProValArgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGlu 760
QY 2865 GTCAATAACGGGTGAGGATATACACCCAGTCCCGAACCTTATACCTTCTGTGT 2924
DB ValAsnLysArgValGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSer 780
QY 2925 GCTCTCCACCCCAACCGAGCTGTATACAGTATTGGACTTAAGGATGCTTCTCTGC 2984
DB AlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCyse 800
QY 2985 CTGAGATTACACCCCACTAGCCAAACCACTTTTTCCTTCCAAATGGAGAGATCCAGGTACG 3044
DB LeuArgLeuHisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThr 820
QY 3045 GGAAGAACGGGCGAGTCACTCGACCCGACTGCCCCAAGGGTTCAGAACTCCCCGACC 3104
DB GlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThr 840
QY 3105 ATCTTTGACGAAGCCCTACACAGACCTCGGCCAACTTCAGGATCCAAACCCCTCAGGTG 3164
DB IlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnVal 860

QY 3165 ACCCTCTCCAGTAGCTGGATGACCTCTTCTGGCGGAGGCCAACACAGACTGCTTA 3224
DB ThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeu 880
QY 3225 GAAGGCACGAAGGCACTACTCTGGAATTTGTCTGACCTAGGCTACAGAGCCTCTGCTAAG 3284
DB GluGlyThrLysAlaLeuLeuLeuLeuSerAspLeuGlyTyrArgAlaSerAlaLys 900
QY 3285 AAGCCCCAGATTTCAGGAGAGAGTAACATACCTTGGGTGTACAGTTTGGCGAGCGGCAG 3344
DB LysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGln 920
QY 3345 CGATGGCTGACGGAGGACCGAAGAAATCTGTAGTCCAGATACCGGCCCCCAACACACGCC 3404
DB ArgTrpLeuThrGluAlaArgLysLysThrValValGlnIleProAlaProThrThrAla 940
QY 3405 AACAAATGAGAGNTTTTGGGACAGCTGGATTTTGCAGACTGTGTGGATCCCGGGTTT 3464
DB LysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPhe 960
QY 3465 GCGACCTTAGCAGCCCACTCTACCCCTAACCAAGAAAAGGGGAATTTCTCTGGGCT 3524
DB AlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLysGlyLysGlyPheSerTrpAla 980
QY 3525 CCTGAGCACGAGGCAATTTGATGTCTATCAAAAAGCCCTGTGTGAGCGCACCTGCTCTG 3584
DB ProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeu 1000
QY 3585 GCCTCTCCTGACGTAACCTTTTACCCCTTTATGTGTGATGAGCTAAGGAGTAGGCC 3644
DB AlaLeuProAspValThrLysProPheThrLeuTyrValAspGluArgLysGlyValAla 1020
QY 3645 CGGGGAGTTTAAACCAACCCCTAGGACCATGAGAGAGACTGTGCGCTCTACCTGTCAAG 3704
DB ArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLys 1040
QY 3705 AAGCTCGATCTGTAGCAGTGTGTGGCCCATATGCTGAAGGCTATCGCAGCTGTGGCC 3764
DB LysLeuAspProValAlaSerGlyTyrProValCysLeuLysAlaIleAlaValAla 1060
QY 3765 ATACTGTCAGACCGCTGACAAATTCAGCTTTGGGACAGAAATTAATCTGTAATAGCCCC 3824
DB IleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaPro 1080
QY 3825 CATGCAATGGAGAAACATCGTTCCGSCAGCCCGCAGACCGATGAGCAACCGCCGCGATG 3884
DB HisAlaLeuGluAsnIleValArgGlnProProAspArgTrpMetThrAsnAlaArgMet 1100
QY 3885 ACCCACTATCAAGCCTGCTTCTCACAGAGGGGTACGTTCCGCTCCACCGCGCTCTC 3944
DB ThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeu 1120
QY 3945 AACCTGCCACTTCTGTGCTGAAGAGCTGATGAACCACTGACTCATGATGCCATCAA 4004
DB AsnProAlaThrLeuLeuProGluThrAspGluProValThrHisAspCysHisGln 1140
QY 4005 CTATTGATTTGAGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAA 4064
DB LeuLeuIleGluGluThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGlu 1160
QY 4065 GTCTAACCTGTTCACTGACGGAAGCAGCTATGTGTGGAAGGTAAAGAGATGGCTGGG 4124
DB ValLeuThrTrpPheThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGly 1180
QY 4125 GCGGCGGTGGTGAACGGGACCCGCAAGATCTGGGCGAGCAGCTGCGGGAAGGAATCCA 4184
DB AlaAlaValValAspGlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSer 1200
QY 4185 GCACAAAGGCTGAGCTCATGCGCTCAGCAGCTTTCGCGCTGCGCGGAGGAATCC 4244
DB AlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSer 1220
QY 4245 ATAAACATTTATACGGACAGCAGGTATGCTTTCGCGACTGCACACGTACATGGGCGCATC 4304

Db 1221 ILeasnlleYrThrAspSerArgTyRAlaPhealThrAlaHisValHisGlyValaile 1240
Qy 4305 TATAAACAAGGGGTGCTTACTCTCAGCAGGGAGGGAATAAAGACAAAGAGAAATT 4364
Db 1241 TyrlYsGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluile 1260
Qy 4365 CTAAAGCCTATTAGAACCGTACATTACCAAAAAGGCTAGCTATTATACATCTGCTCGGA 4424
Db 1261 LeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGly 1280
Qy 4425 CATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATGGCTGACCGGGTTGCCAAG 4484
Db 1281 HisGlnLysAlaLysAspLeuLieserArgGlyAsnGlnMetAlaAspArgValAlaLys 1300
Qy 4485 CAGCAGCCAGGGTGTAACTCTCTGCTTATATAGAAATGCCAAAGCCCCAGAACCC 4544
Db 1301 GlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThrProLysAlaProGluPro 1320
Qy 4545 AGACGACAGTACACCTAGAGACTGGCAAGAGATATAAAGATAGACCACTCTCTGAG 4604
Db 1321 ArgArgGlnTyRThrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGlu 1340
Qy 4605 ACTCCGGAAGGACCTGCTATACCTCAGATGGGAGGAAATCTGCCCCACAAAGAAAGGG 4664
Db 1341 ThrProGluGlyThrCysTyRThrSerTyRgLyLysGluIleLeuProHisLysGluGly 1360
Qy 4665 TTAGAATATATGCCAACAGATACATCGTCTAACCCACCTAGGAACTAAACACCTCGAGAG 4724
Db 1361 LeuGluTyRValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGln 1380
Qy 4725 TTGCTCAGAAATCCCTTATCATGTTCTGAGGCTACGAGAGTGGCTGACTCGGTGGTC 4784
Db 1381 LeuValArgThrSerProTyRHisValLeuArgLeuProGlyValAlaAspSerValVal 1400
Qy 4785 AAACATTGTGCTGCACGCTGTTAATGCTAATCTCTCAGAATGCCCTCCAGGGAAG 4844
Db 1401 LysHisCysValProCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLys 1420
Qy 4845 AGACTAAGGGAAGCCACCCAGGCGCTCACTGGGAAGTGGACTTCACTAGAGTAAAGCCG 4904
Db 1421 ArgLeuArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluValLysPro 1440
Qy 4905 GCTAAATACGAAACAATACTATGCTTTTCTAGACACTTTCAGGATGGGTAGAG 4964
Db 1441 AlalysTyRgLyAsnLysTyRLeuLeuValPheValAspThrPheSerGlyTrpValGlu 1460
Qy 4965 GCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAAAATACTCGAAGAAATT 5024
Db 1461 AlaTyRProThrLysLysGluThrSerThrValValAlaLysLysIleLeuGluGluile 1480
Qy 5025 TTTCCAGATTGGAATACCTAAGTAATAGGTACAGCAATAGTCCAGCTTTTGTGGCC 5084
Db 1481 PheProArgPheGlyIleProLysValIleGlySerAspAsnGlyProAlaPheValAla 1500
Qy 5085 CAGGTAGCTAGGAGCTGCCCAAGATATTGGGGATTGATTGGAACCTGATTGTCATAC 5144
Db 1501 GlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyR 1520
Qy 5145 AGACCCCAAGCTCAGGACAGGTAGAGAGTAAATAGAACCATTAAGACACCTTACT 5204
Db 1521 ArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThr 1540
Qy 5205 AAATGACCGCGGAGACTGGCGTTAATGATTGGATAGCTCTCTGCGCTTTTGTGCTTTT 5264
Db 1541 LysLeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPhe 1560
Qy 5265 AGGTTAGGACACCCCTGGACGTTTGGGCTGACCCCTATGATTAATCTAGCGGGA 5324
Db 1561 ArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyRgLyLeuLeuTyRgLy 1580
Qy 5325 CCCCCCATTTGGTAGAAATTGCTTCTGTACATAGTGTGACGTGCTCTTCCAGCCT 5384

Db 1581 ProProProLeuAlaGluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnPro 1600
Qy 5385 TTGTTCTCTAGGCTCAGGCACCTTGAGTGGGTGAGACAACGAGCGTGGAGCACTCCGG 5444
Db 1601 LeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArg 1620
Qy 5445 GAGGCTTACTCAGGAGGAGGAGACTTCAGATCCACATCGTTTCCAAAGTGGGAGATTCA 5504
Db 1621 GluAlaTyRSer---GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSer 1639
Qy 5505 GTCTACCTTACAGCCACCGTCGAGGAAACCTCGAGACTCGGTGGAAGGGCCCTTATCTC 5564
Db 1640 ValTyRValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyRLeu 1659
Qy 5565 GTACTTTTGACCCACCAACGCGCTGAAAAGTGAAGGAATCTCCACCTGGATCCATGCA 5624
Db 1660 ValLeuLeuThrThrProThrAlaValLysValGluGlyIleProThrTrpIleHisAla 1679
Qy 5625 TCCACACCTTAAACCGCGGCCACCTCCCGATTTCGGGGTGGAAAGCCGAAAGACTGAAAT 5684
Db 1680 SerHisValLysProAlaProProAspSerGlyTrpLysAlaGluLysThrGluAsn 1699
Qy 5685 CCCTTAAGCTTCGCTCCATCGCGTTCCTTACTCTGTCAATAACCTCTCA 5738
Db 1700 ProLeuLysLeuArgLeuHisArgValValProTyRSerValAsnAsnSerSer 1717

RESULT 6
Q5Q08_9GAMR
ID Q5Q08_9GAMR PRELIMINARY; PRT; 1193 AA.
AC Q5Q08;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Pol protein (fragment).
OS Porcine endogenous retrovirus C/A.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammatetrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=286072;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A14/220;
RX PubMed=15564496; DOI=10.1128/JVI.78.24.13880-13890.2004;
RA Bartosch B., Stefanidis D., Myers R., Weiss R., Patience C.,
Takeuchi Y.;
RT "Evidence and consequence of porcine endogenous retrovirus
recombination.";
RL J. Virol. 78:13880-13890 (2004).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY570980; AAT7167.1; -; Genomic DNA.
DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR009007; Pept_Aspatc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR00156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS50879; RNASE H; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW Aspartyl protease; Protease.

```
FT NON_TER 1 1
SQ SEQUENCE 1193 AA; 133310 MW; 9F8EC2AD1A6CE10 CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1193
Score: 6248.50 Matches: 1187
Percent Similarity: 99.7% Conservative: 2
Best Local Similarity: 99.5% Mismatches: 3
Query Match: 42.7% Indels: 1
DB: 2 Gaps: 1

US-10-723-552-3 (1-8132) x Q5QG08_9GMR (1-1193)
QY 2160 GGGAGAGCGGGTTCCGGACCCCTCCCGAGGCCAGGGTAACCTTTGAAGGTGGAGGGCAA 2219
DB 1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGlyGln 20
QY 2220 CCAGTTCAGTTCCTGGTGTATACCGGAGCGAACAATTCAGTGTCTACTACAGCCATTAGGA 2279
DB 21 ProValGluPheLeuValAspThrGlyAlaLysHisSerValLeuLeuGlnProLeuGly 40
QY 2280 AAACTAAAGATAAAAAATCCTGGGTGATGGGTGCCACAGGGCAACACAGTATCCATGG 2339
DB 41 LysLeuLysAspLysLysSerTrpValMetGlyAlaThrGlyGlnGlnGlnTyrSerTrp 60
QY 2340 ACTACCGAAGAACAGTTGACTTCGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCATA 2399
DB 61 ThrThrArgArgThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValIle 80
QY 2400 CCTGAGTCCCGACGACCCCTCTTAGGTAGAGACTATTGACCAAGATGGAGGACAAATT 2459
DB 81 ProGluCysProAlaProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
QY 2460 TCTTTTCAACAGGGAACACAGAGTCTCTGCAATAACAAACCTATCACTGTGTTCGACC 2519
DB 101 SerPheGluGlnGlyLysProGluValSerAlaAsnAsnLysProIleThrValLeuThr 120
QY 2520 CTCCAATTAGATGACGAATATCGACTATCTCTCCCTAGTAAGCCTGATCAAAATATA 2579
DB 121 LeuGlnLeuAspAspGluTyrArgLeuTyrSerProLeuValLysProAspGlnAsnIle 140
QY 2580 CAATTCTGGTTGGAACTGTTCCCAAGCTGGCAGAAACCGCAGGGATGGTTGGCA 2639
DB 141 GlnPheTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuAla 160
QY 2640 AAGCAAGTTCCTCCACCAAGTTATTCAACTGAAGCCAGTCCACACAGGTGTCAAGTCA 2699
DB 161 LysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValArg 180
QY 2700 CAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCCGCATGTCCAAAGATTAAATC 2759
DB 181 GlnTyrProLeuSerLysGluAlaGlnGluGlyIleArgProHisValGlnArgLeuIle 200
QY 2760 CAACAGGGCATCTAGTTCCTGTCCCAATCTCCCTGGNAATCTCCCTGCTACCGGTAGA 2819
DB 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValArg 220
QY 2820 AAGCCTGGGACTAATGACTATCGACAGTACAGGACTTGAGAGAGGTCAATAAACGGTG 2879
DB 221 LysProGlyThrAsnAspTyrArgProValGlnAspLeuArgLysValAsnLysArgVal 240
QY 2880 CAGGATATACACCAACAGTCCCGAACCTTATAACCTCTGTGTGTCTCTCCACCCCAA 2939
DB 241 GlnAspIleHisProThrValProAsnProTyrAsnLeuLeuCysAlaLeuProProGln 260
QY 2940 CGAGCTGGTATACAGTATTGGACTTAAAGATGCTCTTCTCTGCTCGCTGAGATTACACCC 2999
DB 261 ArgSerTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisPro 280
QY 3000 ACTAGCCACCACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAGAACCGGGCAG 3059
DB 281 ThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
```

```
QY 3060 CTCACCTGGACCCCGACTGCCCAAGGGTTCGAAGAACTCCCGACCACTTTTGACGAAGCC 3119
DB LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluAla 320
QY 3120 CTACACAGAGACCTGGCCAACTTCAGGATCCAACACCTCAGGTGAGCCCTCTCCTCAGTAC 3179
DB LeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnTyr 340
QY 3180 GTGATGACCTGCTTCTGGCGGAGCCACCAACACGAGACTGCTTAGAAGGACGACCAAGCA 3239
DB ValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysAla 360
QY 3240 CTACTGCTGAATGTCTGACCTAGGCTACAGAGCTCTCTGCTAAGAAGGCCACAGATTGC 3299
DB LeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCys 380
QY 3300 AGGAGAGAGTAACATACTTGGGTACAGTTTGGGACCGGACCGGACCATGGCTCACGGAG 3359
DB ArgArgGluValThrTyrLeuGlyTyrSerLeuArgAspGlyGlnArgTrpLeuThrGlu 400
QY 3360 GCACGGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACAGCAACAAATCAGAGAG 3419
DB AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGlnMetArgGlu 420
QY 3420 TTTTGGGGACAGCTGGATTTTGCAGACTGTGATCCCGGGTTTGCACCTTTAGCAGCC 3479
DB PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAla 440
QY 3480 CCACTTACCCGCTAACCAAGAAAGGGAATTTCTCTGGGCTCTCTGAGCACCAGAAC 3539
DB ProLeuTyrProLeuThrLysGluLysGlyGluPheSerTrpAlaProGluHisGlnLys 460
QY 3540 GCATTTGATGCTATCAAAAGGCGCTCTGAGCGCACCTCTCTGGGCCCTCCCTGAGCTA 3599
DB AlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspVal 480
QY 3600 ACTAAACCCCTTTTACCCTTTATGTGATGACGCGTAAGGGAGTAGCCCGGGAGTTTAAAC 3659
DB ThrLysProPheThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
QY 3660 CAAACCTTAGGCATGGAGAACCTGTGCGCTACTCTGTCAAAGAAGCTCGATCTCTGTA 3719
DB GlnThrLeuGlyProTrpArgArgProValAlaLysLeuSerLysLysLeuAspProVal 520
QY 3720 GCCAGTGTGTGGCCCATATGCTCTGAAGGCTATCGAGCTGTGCGCATCTACTGTCTCAAG 3779
DB AlaSerGlyTrpProLysLysAlaIleAlaValAlaIleLeuValLysAsp 540
QY 3780 GCTGACAAATTGACTTTGGGACAGAATATACTGTAATAGCCCCCCTGATGGAGAAC 3839
DB AlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluAsn 560
QY 3840 ATCGTTTGGCAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAGC 3899
DB IleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSer 580
QY 3900 CTGCTTCTCACAGAGAGGTTCAGTTCGTCTCCACACGCGCTCTCAACCTGCCACTCTT 3959
DB LeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
QY 3960 CTGCTTGAAGAGACTGATGAACCAAGTCTCATGATTGCCATCACTATTGATTGAGGAG 4019
DB LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluGlu 620
QY 4020 ACTGGGTCGGCAAGACCTTACAGACATACCGCTGACTGGAGAGTGTCTAACTTGGTTC 4079
DB ThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpPhe 640
QY 4080 ACTGACCGAAGCAGCTATGTGTGGAGGTAAAGAGTATGCTGCGGCGCGGTGGTGCAC 4139
DB ThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGlyAlaValValAsp 660
QY 4140 GGGACCCGACGATCTTGGGCCAGCAGCTGCCGGAAGGAATCTTCAGCACAAAAGGCTGAG 4199
```


Db 661 GlyThrArgThrIleTpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaGlu 680
Qy 4200 CTATGCGCCCTCAGCAGAGCTTTGGCTGGCCGCAAGGGAATCCATAAATTTATACG 4259
Db 681 LeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAasnIleTyrThr 700
Qy 4260 GACAGCAGGTATGCTTTGGCAGCTGCACACTGACATGCGGCCATCTATAACAAGGGGG 4319
Db 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgGly 720
Qy 4320 TTGCTTACCTCAGCAGGAGGGAATAAAGCAACAAAGAGGAAATCTTAAGCCTATTAGAA 4379
Db 721 LeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuGlu 740
Qy 4380 GCCTACATTACCAAAAGGCTAGCTATTATACACTGCTCTCTGCACATCAGAAAGCTAAA 4439
Db 741 AlaValHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaLys 760
Qy 4440 GATCTCATATCCAGAGNAACCATGCTGACCGGTTGCCAAGCAGGAGCCAGGGT 4499
Db 761 AspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaAsnGlnGly 780
Qy 4500 GTTAACTTCTGCTTATAATAGAAATGCCAAAGCCCCAGAACCCAGACGACAGTACACC 4559
Db 781 ValAsnLeuLeuProIleIleGluMetProLysAlaProGluProArgGlnTyrThr 800
Qy 4560 CTAGAAGACTGGCAAGAGATAAAAGATAGACAGACTTCTCTGAGACTCCGGAGGGACC 4619
Db 801 LeuGluAspTyrGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyThr 820
Qy 4620 TGCTATACCTCAGATGGGAAGAAATCTGCCCCACAAAGAGGGTTAGATATGTCCAA 4679
Db 821 CysTyrThrSerAspGlyLysGluIleLeuProHisLysGluGlyLeuGluTyrValGln 840
Qy 4680 CAGATACATCTCTAAACCCACTAGGAACCTAAACACCTGCAGCAGTGGTGGTCAGAACATCC 4739
Db 841 GlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrSer 860
Qy 4740 CCTATCATGTTCTGAGCTACAGGAGTGCTGACTCGGTGGTCAACATTTGTGTCCT 4799
Db 861 ProTyrHisValLeuArgLeuProGlyValAlaAspSerValLysHisCysValPro 880
Qy 4800 TGCCAGCTGGTTATGCTAATCCCTCCAGATGCTCCAGGGAGAGACTAAGGGAGC 4859
Db 881 CysGlnLeuValAsnAlaAsnProSerArgMetProProGlyLysArgLeuArgGlySer 900
Qy 4860 CACCAGCGCTCACTGGGAAGTGGACTTCACTCAGGTAAAGCCGGCTAAATACGGAAC 4919
Db 901 HisProGlyAlaHisTyrGluValAspPheThrGluValLysProAlaLysTyrGlyAsn 920
Qy 4920 AAATACCTATTGTTTGTAGACACCTTTTTCAGATGGGTAGAGGCTTATCTCTACTAAG 4979
Db 921 LysTyrLeuLeuValPheValAspThrPheSerGlyTyrValGluAlaTyrProThrLys 940
Qy 4980 AAAGAGACTTCAACCGTGGCTTAAATAAATACTGGAGAAATTTTCCAGATTGGA 5039
Db 941 LysGluThrSerThrValValAlaLysLysIleLeuGluGluIlePheProArgPheGly 960
Qy 5040 ATACTAAGGTAATAGGTCAGACAATGGTCCAGCTTTTGTGTCAGGTAAAGTCAGGGA 5099
Db 961 IleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnGly 980
Qy 5100 CTGCGCAAGATATTGGGATTTGATTGGAACTGCAATTTGTGCATACAGACCCCAAGCTCA 5159
Db 981 LeuAlaLysIleLeuGlyIleAspTyrLysLeuHisCysAlaTyrArgProGlnSerSer 1000
Qy 5160 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGACCTTACTTAATTCACCGCGAG 5219
Db 1001 GlyGlnValGluArgMeCAsnArgThrIleLysGluThrLeuThrLysLeuThrAlaGlu 1020
Qy 5220 ACTGGCGTTAATGATTGGATAGCTCTCCTCGCCCTTTGTGCTTTTGTAGGGTTAGGAACACC 5279

Db 1021 ThrGlyValAsnAspTyrIleAlaLeuLeuProPheValLeuPheArgValArgAsnThr 1040
Qy 5280 CCTGGACAGTTTGGCTGACCCCTATGAATTACTCTACGGGGACCCCCCAATTTGTA 5339
Db 1041 ProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrGlyGlyProProLeuVal 1060
Qy 5340 GAAATCTCTCTGTACATAGTGTGCTGCTTTCCAGCCCTTTGTTCTCTAGGCTC 5399
Db 1061 GluIleAlaSerValHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeu 1080
Qy 5400 AAGCATTGAGTGGTGAGAACAGCGTGGAGCAACTCCGGAGGCCCTACTCAGGA 5459
Db 1081 LysAlaLeuGluTyrValArgGlnArgAlaTyrLysGlnLeuArgGluAlaTyrSer 1099
Qy 5460 GGAGGACACTTGCAGATCCCATCTGTTCCAGTGGGAGATTTCAGTCTAGTTAGCGC 5519
Db 1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrValArg 1119
Qy 5520 CACCGTCAGCAACCTCGAGACTCGGTGGAGGCCCTTATCTCTGACTTTTTCACCACA 5579
Db 1120 HisArgAlaGlyAsnLeuGluThrArgTyrLysGlyProTyrLeuValLeuLeuThrThr 1139
Qy 5580 CCAACGCTGTGAAGTCGAAGGAATCTCCACCTGGATCCATGATCCCATGTTAAACCG 5639
Db 1140 ProThrAlaValLysValGluGlyIleProThrTyrIleHisAlaSerHisValLysPro 1159
Qy 5640 GCGCACCTCCGATTCGGGGTGGAAAGCCGAAAGACTGAAATCCCTTAAAGCTTCGC 5699
Db 1160 AlaProProAspSerGlyTyrLysAlaGluLysThrGluAsnProLeuLysLeuArg 1179
Qy 5700 CTCATCGCTGGTTCCTTACTCTGTCAATAACCTCTCA 5738
Db 1180 LeuHisArgValValProTyrSerValAsnAsnSerSer 1192

RESULT 7
Q9ORL9_9GAMR
ID Q9ORL9_9GAMR PRELIMINARY; PRT; 1193 AA.
AC Q9ORL9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Name=pol;
OS Porcine endogenous type C retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gamma retrovirus.
OX NCBI_TaxID=70540;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21256017; PubMed=11356953;
RX DOI=10.1128/JVI.75.12.5465-5472.2001;
RA Krach U., Fischer N., Czaderna F., Toenjes R.R.;
RT "Comparison of replication-competent molecular clones of porcine
endogenous retrovirus class A and class B derived from pig and human
cells.";
RL J. Virol. 75:5465-5472(2001).
CC - SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AJ293656; CAC39617.1; -; Genomic_DNA.
DR HSSP; P03355; 116J.
DR MEROPS; A02.020; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; F:DNA recombination; IEA.
DR GO; GO:0006508; P:proteinolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR009007; Pept_Asparts_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.

DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS50879; RNASE H; 1.
DR PROSITE; PS50878; RT POL; 1.
KW Aspartyl protease; Protease.
FT NON TRS
SQ SEQUENCE 1193 AA; 133225 MW; 8FC1F419CB600332 CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1193
Score: 6170.50 Matches: 1168
Percent Similarity: 98.9% Conservative: 12
Best Local Similarity: 97.9% Mismatches: 12
Query Match: 42.2% Indels: 1
DB: 2 Gaps: 1

US-10-723-552-3 (1-8132) x Q90RL9_9GMR (1-1193)

QY 2160 GGGAGAGCGGGTTTCGGACCCCTCCCGAGCCCGAGGTAACTTTGAAGGTGGAGGGCAA 2219
DB 1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGlyGln 20

QY 2220 CCAGTTCAGTTCCTGGTTGATACCGGAGCGGAAACATTTCAGTGTCTACTACAGCCATTAGGA 2279
DB 21 ProValGluPheLeuValAspThrGlyAlaLysHisSerValLeuLeuGlnProLeuGly 40

QY 2280 AAACTAAAAGATAAAAAATCTCGGTGATGGTGCCACAGGGCAAACAACAGTATCCATGG 2339
DB 41 LysLeuLysAspLysLysSerTrpValMetGlyAlaThrGlyGlnGlnGlnTyPProTrp 60

QY 2340 ACTACCCGAAGAACAGTGTGACTTCGGAGTGGAGCGGTAAACCCACTCGTTCTCGTGCATA 2399
DB 61 ThrThrArgArgThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValIle 80

QY 2400 CCTGAGTCCCGACACCCCTTAGGTAGAGACTTATTGACCAAGATGGGAGCAAAATT 2459
DB 81 ProGluCysProAlaProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100

QY 2460 TCTTTTCAACAGGGAACAGAGTGTCTGCAATAACAAACCTATCACTGTGTGAC 2519
DB 101 SerPheGluGlnGlyLysProGluValSerAlaAsnAsnLysProIleThrValLeuThr 120

QY 2520 CTCGAATTAGATGACGAATATCGACTATCTCTCCCTAGTAAAGCCCTGATCAAAATATA 2579
DB 121 LeuGlnLeuAspAspGluTyArgLeuTySerProLeuValLysProAspGlnAsnIle 140

QY 2580 CAATTCTGGTTGNAAGTTCCTCCAGCTGGCAGAAACCGGAGGATGGGTTGGCA 2639
DB 141 GlnPheTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuAla 160

QY 2640 AAGCAAGTCTCCCCACAGTATTTCACCTGAAGCCAGTGCCACACAGGTGTCAGTCAGA 2699
DB 161 LysGlnValProProGlnValIleGlnLeuLysAlaSerAlaProValSerValArg 180

QY 2700 CAGTACCCCTTGAGTAAAGAGCTCAAGAAGGAATTCGGCCGATGTCCAAAGATTATC 2759
DB 181 GlnTyProLeuSerLysGluAlaArgGluGlyIleArgProHisValGlnArgLeuIle 200

QY 2760 CAACAGGGATCCTAGTTCCTGTCCCAATCTCCCTGGAAATACTCCCTGTCTACCGGTTAGA 2819
DB 201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValArg 220

QY 2820 AAGCCTGGGACTAATGACTATCGACCAAGTACAGGACTTGAGAGAGGTCAATAACGGGTG 2879
DB 221 LysProGlyThrAsnAspTyArgProValGlnAspLeuArgGluValAsnLysArgVal 240

QY 2880 CAGGATATACACCAACAGTCCCGAACCCCTTATAACCTCTTGTGTCTCTCCACCCCAA 2939

DB 241 GlnAspIleHisProThrValProAsnProTyArgAsnLeuLeuCysAlaLeuProGln 260
QY 2940 CGGAGCTGTATACAGTATTGGACTTAAAGATCCCTCTTCTGCTGAGATTACACCCC 2999
DB 261 ArgSerTrpTyThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisPro 280
QY 3000 ACTAGCCAAACACATTTTTCCTCCAAATGGAGAGATCCAGTACGGGAAGAACCGGGCAG 3059
DB 281 ThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyAlaGlyArgThrGlyGln 300
QY 3060 CTCACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCGACCATCTTTTGACGAAGCC 3119
DB 301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluAla 320
QY 3120 CTACACAGAGACCTGGCCAACTTCAGGATCCAAACACCTCAGGTAGCCCTCTCCAGTAC 3179
DB 321 LeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnTy 340
QY 3180 GTGATGACCTGTCTTCGGCGGAGCCACCAACAGGACTGTCTAGAGGACCAAGCA 3239
DB 341 ValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysAla 360
QY 3240 CTACTGCTGAATTGTCTGACCTAGGCTACAGACCTCTGCTAAGAGGCCCCAGATTGC 3299
DB 361 LeuLeuLeuGluLeuSerAspLeuGlyTyArgAlaSerAlaLysAlaGlnIleCys 380
QY 3300 AGGAGAGAGTAACATCTTGGGTACAGTTTCGGGACCGGCGAGCCATGGCTGACGGAG 3359
DB 381 ArgArgGluValThrTyLeuGlyTySerLeuArgGlyGlnArgTrpLeuThrGlu 400
QY 3360 GCACGGGAAGAAACTGTAGTCCAGATACCGGCCCAACACAGCAACCAAAATCAGAGAG 3419
DB 401 AlaArgLysArgThrValValGlnIleProAlaProThrThrAlaLysGlnValArgGlu 420
QY 3420 TTTTGGGGACAGCTGGATTTCAGACTGTGATCCCGGGTTTGGACCTTTAGCAGCC 3479
DB 421 PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAla 440
QY 3480 CCATCTACCCGCTAACCAAGAAAAGGGAATTCCTCGGCTCTCTGAGCACCAGAAC 3539
DB 441 ProLeuTyProLeuThrLysGluLysGlyGluPheSerTrpAlaProGluHisGlnLys 460
QY 3540 GCATTTGATCTATCAAAAAGGCCCTGTGAGCGCACCTCTCTGGGCCCTCCCTGAGCTA 3599
DB 461 AlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspVal 480
QY 3600 ACTAAACCCCTTTACCCCTTTATGTGGATGAGCGTAGCGGAGTAGCCCGGGAGTTTAAAC 3659
DB 481 ThrLysProPheThrLeuTyValAspGluArgLysGlyValAlaArgGlyValLeuThr 500
QY 3660 CAACCCCTAGGACCATGGAGAACCTGTGCGCTACTCTCAAGAAGACTCGATCTCTGTA 3719
DB 501 GlnThrLeuGlyProTrpArgProValAlaTyLeuSerLysLysLeuAspProVal 520
QY 3720 GCCAGTGTGGGCCCATATGCTGAAGGCTATCGAGCTGTGCGCATCTAGTGTCAAGGAC 3779
DB 521 AlaserGlyTrpProLysLeuLysAlaIleAlaValAlaIleLeuValLysAsp 540
QY 3780 GCTGACAAATTCATTTGGGACAGAAATATAACTGTAAATAGCCCCCATGATCTGGAGAAC 3839
DB 541 AlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluAsn 560
QY 3840 ATGCTTGGGAGCCCGGACCGATGATGACCAACCGCCCGCATGACCCCACTATCAAGC 3899
DB 561 IleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyArgGlnSer 580
QY 3900 CTGCTTCTCACAGAGAGGTACGTTTCGTCTCCACAGCGCTCTCAACCTGCCACTCTT 3959
DB 581 LeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
QY 3960 CTGCTTGAAGAGACTGATGAACACAGTGACTCATGATTGGCATCACTATTGATTGAGGAG 4019

Db 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuLeuGluGlu 620
Qy 4020 ACTGGGTCGCCAAGAGCCTTACAGACATACCGCTGACTGGAGAAGTGCCTAACCTGGTTC 4079
Db 621 ThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluMetLeuThrTrpPhe 640
Qy 4080 ACTCAGCGAAGCAGCATGTGCTGGAAGTGAAGAGGATGGCTGGGGCGGCGTGGTGGAC 4139
Db 641 ThrAspGlySerSerTrpMetValGluGlyLysArgMetAlaGlyAlaIvalValAsp 660
Qy 4140 GGGACCCGACGATCTGGCCAGCAGCTGCCGAAGAACTTCAGACACAAAAGCTGAG 4199
Db 661 GlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaGlu 680
Qy 4200 CTGATGGCCCTCAGCAAGCTTTGGGCTGGCCGAAGGAAATCCATAAACAATTTATACG 4259
Db 681 LeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTrpThr 700
Qy 4260 GACAGCAGGTATGCTTTGGCACTGCACACGTACATGGGGCATCTATAACAAGGGGG 4319
Db 701 AspSerArgTrpAlaPheAlaThrAlaHisValHisGlyAlaIleTrpLysGlnArgGly 720
Qy 4320 TTGCTTACCTCAGCAGGGAGGAAATAAGAACAAAGAGAAATCTTAAGCCTATTAGAA 4379
Db 721 LeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuGlu 740
Qy 4380 GCCGTACATTTACAAAAGGCTAGCTATTATATACACTGCTCTGGACATCAGAAAGCTTAA 4439
Db 741 AlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaLys 760
Qy 4440 CATCTATATCCAGAGAAACACAGATGCTCACCGGTTGCTCAAGCAGCGACGCCAGGT 4499
Db 761 AspProIleSerArgLysGlnMetAlaAspArgValAlaLysGlnAlaIleGlnGly 780
Qy 4500 GTTAACCTTCTGCTTATATAAGAAATGCCAAAGCCCCAGAACCCAGACGACGATACAC 4559
Db 781 ValAsnLeuLeuProMetIleGluThrProLysAlaProGluProGlyArgIleTrpThr 800
Qy 4560 CTAGAAGACTGGCAAGAGATAAAAGATAGACACAGTTCTCTGAGACTCCCGAAGGGACC 4619
Db 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyThr 820
Qy 4620 TGCTATACCTCAGATGGGAAGAAATCCTGCCCAAGAAAGGTTAGATATATGTCCA 4679
Db 821 CysTrpThrSerAspGlyLysGluIleLeuProHisLysGluGlyLeuGluTrpValGln 840
Qy 4680 CAGATACATCTCTAACCCACTAGGAACCTGACAGCTGGTGGTGGTGGTGGTGGTGGT 4739
Db 841 GlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrSer 860
Qy 4740 CTTATCATGTTCTGAGGCTACCAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4799
Db 861 ProTrpHisValLeuArgLeuProGlyValAlaAspSerValValLysHisCysValPro 880
Qy 4800 TGCCAGCTGGTTAATGCTAATCTCTCCAGAATGCTCCAGGGAAGAGACTAAGGGGAGC 4859
Db 881 CysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLysArgLeuGlySer 900
Qy 4860 CACCAGCGCTCACTGGGAAGTGGACTTCACTCAGAGTAAAGCCGGCTAAATACGGAAC 4919
Db 901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLysTrpGlyAsn 920
Qy 4920 AAATACCTATTGTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCTCTACTAAG 4979
Db 921 LysTrpLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaLysProThrLys 940
Qy 4980 AAGAGACTTCAACCGTGGTCTTAAATAAATACTCGAAGAAATTTTCCAGATTTTGA 5039
Db 941 LysGluThrSerThrValValAlaLysLysIleLeuGluGluIlePheProArgPheGly 960
Qy 5040 ATACCTAAGGTAATAGGCTCAGACATGCTCCAGCTTTTGTGGCCAGGTAAAGTCAGGA 5099
Db 961 IleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnGly 980

Qy 5100 CTGCCAAGATATTGGGGATTGATTGGAACCTGATTGTGCATACAGACCCCAAGCTCA 5159
Db 981 LeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTrpArgProGlnSerSer 1000
Qy 5160 GGACAGCTAGAGAGGATGAATAGAACCATTAAGAGACCCCTTACTATAAATTGACCGCGAG 5219
Db 1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrGlu 1020
Qy 5220 ACTGGCGTTAATGATTGGATAGCTCTCTCCCTTGTGCTTTTGTAGGGTTAGGAAACAC 5279
Db 1021 ThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnThr 1040
Qy 5280 CCTGACAGTTTGGCTGACCCCTATGAATTAATCTACGGGGACCCCGCCCATTTGGTA 5339
Db 1041 ProGlyGlnPheGlyLeuThrProTrpGluLeuLeuTrpGlyGlyProProLeuAla 1060
Qy 5340 GAAATCTTCTGTACATAGTGTGCTGCTGCTTCCAGCGCTTTTCTCTAGGCTC 5399
Db 1061 GluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeu 1080
Qy 5400 AAGCAGCTTGGTGGTGAGACAGCGTGGAGGCAACTCCGGGAGGCTACTACGGA 5459
Db 1081 LysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTrpSer 1099
Qy 5460 GGAGGAGCTTGCAGATCCACATCGTTTCCAGTGGGAGATTCACTAGTTAGACGC 5519
Db 1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTrpValArg 1119
Qy 5520 CACGCTCAGAGAAACCTCGAGACTCGGTGGAGGGCCCTTATCTCGTACTTTTACCACA 5579
Db 1120 HisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTrpLeuValLeuThrThr 1139
Qy 5580 CCAACGGCTGTGAAGTCTGAAGGATCTCCACCTGGATCCATGCTATCCAGCTTAAACCG 5639
Db 1140 ProThrAlaValLysValGluGlyIleProThrTrpIleHisAlaSerHisValBPro 1159
Qy 5640 GCGCACCTCCGATTCGGGTGGAAAGCCGAAAGACTGAAATCCCTTAAAGCTTCG 5699
Db 1160 AlaProProAspSerGlyTrpLysAlaGluLysThrGluAsnProLeuLysLeuArg 1179
Qy 5700 CTCCATCGCTGGTTCCTTACTCTGTCAATAACCTCTCA 5738
Db 1180 LeuHisArgValValProTrpSerValAsnAsnSerSer 1192
RESULT 8
Q90RL6_9GAMR PRELIMINARY; PRT; 1195 AA.
AC Q90RL6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Name=pol;
OS Porcine endogenous type C retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammaretrovirus.
OX NCBI_taxid=70540;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=21256017; PubMed=11356953;
RX DOI=10.1128/JVI.75.12.5465-5472.2001;
RA Krach U., Fischer N., Czaderna F., Toenjes R.R.;
RT "Comparison of replication-competent molecular clones of porcine
RT endogenous retrovirus class A and class B derived from pig and human
RT cells.";
RL J. Virol. 75:5465-5472(2001).
CC -I- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AJ293657; CAC39626.1; -; Genomic_DNA.
DR HSSP; P03355; 116J.
DR MEROPS; A02.020; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR InterPro; IPR009007; Pept_Aspartc_cat.
 DR InterPro; IPR001995; Peptidase_A2_cat.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR012337; RNaseH_fold.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVtse.
 DR Pfam; PF00075; RNaseH.1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT.1; 1.
 DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS00994; INTEGRASE; 1.
 DR PROSITE; PS00879; RNASE H; 1.
 DR PROSITE; PS00878; RT_POL; 1.
 DR KW Aspartyl protease; Hydrolase; Protease.
 FT NON_TER 1
 SQ SEQUENCE 1195 AA; 133770 MW; 999F1524D50BFFA4 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 1195
 Score: 6109.00 Matches: 1155
 Percent Similarity: 98.6% Conservative: 19
 Best Local Similarity: 97.0% Mismatches: 17
 Query Match: 41.7% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3 (1-8132) x Q09RL6_9GMR (1-1195)

QY	2160	GGGAGACGGGTTCCGACCCCTCCCGAGCCAGGGTAACTTTGAAGGTGGAGGGCAA	2219		
DB	1	GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGln	20		
QY	2220	CCAGTTGACCTCTGGTTGATACCGGAGCGAAACATTCAGTGTCTACTACAGCCATTAGGA	2279		
DB	21	ProValGluPheLeuValAspThrGlyAlaGluHisSerValLeuLeuGlnProLeuGly	40		
QY	2280	AAACTAAAGATAAAATCCTGGTGCATGGTCCACAGGGCAACACAGTATCCATGG	2339		
DB	41	LysLeuLysGluLysLysSerTrpValMetGlyAlaThrGlyGlnArgGlnTrpProTrp	60		
QY	2340	ACTACCCGAAGAACAGTTGACTTGGGAGTGGGAGTAAACCCACTCGTTTCTGGTCATA	2399		
DB	61	ThrThrArgArgThrValAspLeuAlaValGlyArgValThrHisSerPheLeuValIle	80		
QY	2400	CCTGAGTGGCCAGCACCCCTCTTAGGTAGAGACTTATGACCAAGATGGAGGACAAATT	2459		
DB	81	ProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle	100		
QY	2460	TCTTTTCAACAGGGAACAGAAAGTCTCTGCAATAACAACTATCATCTGTGTGAC	2519		
DB	101	SerPheGluGlnArgProGluValSerValAsnAsnLysProIleThrValLeuThr	120		
QY	2520	CTCCAAATTAGATGACGAATATCGACTATCTCTCCCTAGTAAAGCCGTGATCAAAATATA	2579		
DB	121	LeuGlnLeuAspAspGluTyArgLeuTySerProGlnValLysProAspGlnAspIle	140		
QY	2580	CAATTCGTGGTGAACAGTTTCCCAAGCTGGCGAGAAACCGAGGATGGGTTGGCA	2639		
DB	141	GlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuAla	160		
QY	2640	AACCAAGTCCCCACAGTATTCACTCAAGCCAGTGCACACCCAGTGTCAAGTCA	2699		
DB	161	LysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValArg	180		
QY	2700	CAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCGGCATGTCCAAAGATTAATC	2759		
DB	181	GlnTyProLeuSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle	200		
QY	2760	CAACAGGGCATCCTAGTTCCTGTCCTCAATCTCCCTGGAAATCTCCCTGCTACCCGTTAGA	2819		
DB	201	GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValArg	220		
QY	2820	AAGCCTCGGACTAATGACTATCGACCACTACAGGACTTGAGAGAGGTCAATAAAGCGGTG	2879		
DB	221	LysProGlyThrAsnAspTyArgProValGlnAspLeuArgGluValAsnLysArgVal	240		
QY	2880	CAGGATATACACCCACAGTCCCGAACCTTATAACCTTCTGTGTGCTCTCCACCCCAA	2939		
DB	241	GlnAspIleHisProThrValProAsnProTyAsnLeuLeuSerAlaLeuProProGlu	260		
QY	2940	CGGAGCTGGTATACAGTATTGGACTTAAAGATGCCTTCTTCCTCGCTGAGATTACACCC	2999		
DB	261	ArgAsnTrpTyThrValLeuAspLeuLysAspAlaPheCysLeuArgLeuHisPro	280		
QY	3000	ACTAGCCCAACACTTTTTCCTTCGAATCGAGAGATCCAGTACGGGAAGAACCCGGCAG	3059		
DB	281	ThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln	300		
QY	3060	CTCACCTGGACCCGACTGCCCCAAGGTTCAAGAATCTCCCGACCATCTTTTGACGAAGCC	3119		
DB	301	LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrPheAspGluAla	320		
QY	3120	CTACACAGACACTGGCCAACTTCAGGATCCACACACCTCAGGTGACCTCTCCAGTAC	3179		
DB	321	LeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnTy	340		
QY	3180	GTGGATGACCTGCTCTGCGCGGAGCCACCAACAGGACTCTTAGAAGGCACGAAGGCA	3239		
DB	341	ValAspAspLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysAla	360		
QY	3240	CTACTCTGGAATGTCTGACCTAGGCTACAGAGCCTCTCTAAGAGGCCAGATTTC	3299		
DB	361	LeuLeuLeuGluLeuSerAspLeuGlyTyArgAlaSerAlaLysLysAlaGlnIleCys	380		
QY	3300	AGGAGAGAGTAACACTTGGGTACAGTTTCGGGACGGGACGGAGCGATGGTACCGAG	3359		
DB	381	ArgArgGluValThrTyLeuGlyTySerLeuArgGlyGlyGlnArgTrpLeuThrGlu	400		
QY	3360	GCACGAAGAAAACCTGTAGTCCAGATACCGGCCCAACACACGCAAAACAAATAGAGAG	3419		
DB	401	AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGlnValArgGlu	420		
QY	3420	TTTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGTTTTCGACCTTAGCAGCC	3479		
DB	421	PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAla	440		
QY	3480	CGACTCTACCGGTAAACCAAGAAAAGGGGAATTCCTCTGGGCTCTCTGAGCACCAGAAG	3539		
DB	441	ProLeuTyProLeuThrLysGluLysGlyGlyPheSerTrpAlaProGluHisGlnLys	460		
QY	3540	GCATTTGATGCTATCAAAAGGCCCTGTGAGCGCACCTGTCTGGCCCTCCCTGACGTA	3599		
DB	461	AlaPheAspAlaLysLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspVal	480		
QY	3600	ACTAAACCCCTTTTACCCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGGGGAGTTTAAAC	3659		
DB	481	ThrLysProPheThrLeuTyValAspGluArgLysGlyValAlaArgGlyValLeuThr	500		
QY	3660	CAAAACCTTAGACCATGGAGAGACCTGTGCGCTACTGTCTGCTCAAGAGCTCGATCTGTGA	3719		
DB	501	GlnThrLeuGlyProTrpArgArgProValAlaTyLeuSerLysLysLeuAspProVal	520		
QY	3720	GCCAGTGGTGGCCCATATGCTGAAGGCTATCGCAGCTGTGCGCATCTGTCAGTCAGGAC	3779		
DB	521	AlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLeuValLysAsp	540		
QY	3780	GCTGACAAATTGACTTTGGGACAGAAATATAACTGTAAATAGCCCCCATGCAATTGGAGAAC	3839		

RC STRAIN=minature swine;
RX MEDLINE=98216827; PubMed=9557749;
RA Akiyoshi D.E., Denaro M., Zhu H., Greenstein J.L., Banerjee P.,
RA Fishman J.A.;
RT "Identification of a full-length cDNA for an endogenous retrovirus of
RT miniature swine.";
RL J. Virol. 72:4503-4507(1998).
DR EMBL; AF038600; AAC16764.1; -; mRNA.
DR HSSP; P03355; 116J.
DR MEROPS; A02.020; -.
DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; RVP; 1.
DR PROSITE; PS01078; RVT_1; 1.
DR PROSITE; PS0175; ASP PROT RETROV; 1.
DR PROSITE; PS0879; RNASE H; 1.
SQ SEQUENCE 1145 AA; 128010 MW; 90C1A95E9D95B88 CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1145
Score: 6026.00 Matches: 1143
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 2
Query Match: 41.2% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3 (1-8132) x Q9XS88_PIG (1-1145)

QY 2307 ATGGTGCCACGCGGCAACACAGTATCCATGGACTACCCGAGACAGTGTGACTGGGA 2366
DB 1 MetGlyAlaThrGlyGlnGlnTyrProTrpThrThrArgArgThrValAspLeuGly 20

QY 2367 GTGGGACGGTAAACCCACTCGTTCTCGTCTATCTGAGTCCAGCCAGCCCTCTTAGGT 2426
DB 21 ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly 40

QY 2427 AGAGACTTATTGACCAAGATGGGAGCAAAATTTCTTTTGAACAGGGAACAGAGTG 2486
DB 41 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal 60

QY 2487 TCTGCAATTAACAACTATCAGTGTGACCTCCAAATAGATGACGAATATCGACTA 2546
DB 61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 80

QY 2547 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATTTCTGTTGGAACAGTTTCCCAA 2606
DB 81 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGlnPheProGln 100

QY 2607 GCTGGGACAGAACCCGAGGATGGGTTTGGCAAGCAAGTCCCCCAAGTTATTCAA 2666
DB 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120

QY 2667 CTGAGGCCAGTGCACACAGTGTCTAGTCAGACAGTACCCTTGAGTAAAGAGCTCAA 2726
DB 121 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln 140

QY 2727 GAAGGAATTCGGCGGATGTCCAAAGATTAAATCCACAGGCGCATCTAGTTCCTGTCCAA 2786
DB 141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 160

QY 2787 TCTCCCTGGGAATACTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCA 2846

DB 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
QY 2847 GTACAGACTTGTAGAGAGGTCAATAAACGGGTGCAGGATATACACCCACAGTCCCGAAC 2906
DB 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200
QY 2907 CCTTATAACCTCTGTGTCTCTCCACCCCAACGGAGCTGTATACAGTATTGACACTTA 2966
DB 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpThrValLeuAspLeu 220
QY 2967 AAGGATCCCTTCTTCTGCTGAGATTACACCCACTAGCCCAACACTTTTTCCTTCGAA 3026
DB 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
QY 3027 TGGAGAGATCCAGGTACGGAGAACCGGACGCTCACTCGGACCCAGCTGCCCCNAGG 3086
DB 241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
QY 3087 TTCAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTCAGG 3146
DB 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280
QY 3147 ATCCAACACCTCAGGTGACCTCTCTCCAGTACGTGGATCAGCTCTTCTGGCGGAGCC 3206
DB 281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 300
QY 3207 ACCAAACAGACTGCTTAGAAGGACGACGACACTACTGCTGGNATTGTCTGACCTAGG 3266
DB 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 320
QY 3267 TACAGAGCTCTGCTAAGAGGCCACGATTTGACGAGAGAGGTAAACATCTTGGGGTAC 3326
DB 321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 340
QY 3327 AGTTTGGGACCGGACGCGATCGCTGACGAGGACGACGAGAAACCTGTAGTCCAGATA 3386
DB 341 SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 360
QY 3387 CCGGCCCCACACACGACCAACAACATGAGAGAGTGTGGGACAGCTGGATTTGCAGA 3446
DB 361 ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg 380
QY 3447 CTGTGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAGAAA 3506
DB 381 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 400
QY 3507 GGGGAATTCCTCTGGGCTCCTGAGCACCAAGAGGCAATTTGATGCTATCAAAAAGGCCCTG 3566
DB 401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
QY 3567 CTGAGGCGACTGCTCTGGCCCTCCCTGAGCTAACTAAACCTTTTACCTTTATGTGGAT 3626
DB 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440
QY 3627 GACGCTAAGGAGTACCGGGGAGTTTAAACCAACCCCTAGGACCATCGAGAGACCT 3686
DB 441 GluArgLysGlyValAlaAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 460
QY 3687 GTCGCTACTCTGTCAAAGAGCTCGATCTCTGTAGCCAGCTGTTGGGCCATATGCTCAAG 3746
DB 461 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLys 480
QY 3747 GCTATCGAGCTGTGGCCATCTAGTCAAGGACGCTGACAAATGTACTTTGGGACAGAAAT 3806
DB 481 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500
QY 3807 ATAACGTATAGCCCCCATGTCATTGGAGAACATCGTTCCGCGAGCCCCCAGACCCGATGG 3866
DB 501 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 520
QY 3867 ATGACCAACGCCCGCATGACCCACTATCAAGCCTCTCTCTCACAGAGGGGTCAAGTTC 3926

Db 521 MetThrAsnAlaArgMetThrHisThrGlnSerLeuLeuLeuLeuThrGluArgValThrPhe 540
Qy 3927 GCTCACCAGCGCTCTCAACCCCTGCCACTCTTCTGCTGAAAGACTGATGAACCAAGTG 3986
Db 541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuLeuProGluGluThrAspGluProVal 560
Qy 3987 ACTCATGATTCCTCCATCAACTATTGATTGAGGAGACTGGGGTCCCAAGGACTTTACAGAC 4046
Db 561 ThrHisAspCysHisGlnLeuLeuLeuLeuGluGluThrGlyValArgLysAspLeuThrAsp 580
Qy 4047 ATACCGCTGACTGGAGAGTGCTAACCTGGTTCTACGACGGAACAGACTATGTGGTGAA 4106
Db 581 IleProLeuThrGlyGluValLeuLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
Qy 4107 GGTAAAGAGATGGCTGGGGCGGGTGTGACGGGACCCGACGATCTGGCCAGCAGC 4166
Db 601 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
Qy 4167 CTGCCGGAAGAACTTCAGACACAAAGGCTGAGCTCATGGCCCTCAGCGAGCTTTGCGG 4226
Db 621 LeuProGlyGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuLeuThrGlnAlaLeuArg 640
Qy 4227 CTGCCGGAAGGAAATCCATAAAACATTTATACGACAGCAGGTATGCCCTTTGCGACTGCA 4286
Db 641 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
Qy 4287 CACGTACATGGGGCATCTATAACAAAGGGGTTGCTTACCTCAGCAGGGAGGGAATA 4346
Db 661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
Qy 4347 AAGAACAAAGAGGAATTTCTAAGCCTATTAGAAGCCGTACATTTACCAGAAAGGCTAGCT 4406
Db 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
Qy 4407 ATTATACACTCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACCATG 4466
Db 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
Qy 4467 GCTCAGCGGTTGCCAAGCAGCGACCCAGGGTGTAACTTCTGCTTATTAATAGAATG 4526
Db 721 AlaAspArgValAlaLysGlnAlaGlnGlyValAsnLeuLeuProIleIleGluMet 740
Qy 4527 CCCAAAGCCACAGAACCCAGACGACGATACACCTTAGAAGACTGCAAGACATATAAAG 4586
Db 741 ProLysAlaProGluProArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
Qy 4587 ATAGACCAAGTTCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATC 4646
Db 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
Qy 4647 CTGCCCCCAAGAGGGTTAGAATATGTCACACAGATACATGCTTAACCCACTAGGA 4706
Db 781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
Qy 4707 ACTAAACACCTCGACGCTGTGACAGACATCCCTTATCATGTTCTGAGGCTACCAAG 4766
Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
Qy 4767 GTGGCTGACTCGGTGGTCAAACTTGTGTGCCCTGCCAGCTGGTTAATGCTTAATCTTCC 4826
Db 821 VallalaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
Qy 4827 AGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCACTGGGAAGTGGAC 4886
Db 841 ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
Qy 4887 TTCACAGGTAAAGCCGCTTAATACCGGAACAAATACCTATTGGTTTTGTAGACACC 4946
Db 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
Qy 4947 TTTTCAGGATGGGTAGAGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 5006
Db 881 PheSerGlyTyrValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900

Qy 5007 ABAATACCTGAAGAAATTTTTCGAAGATTGGAAATACCTAAGGTAATAAGGTCAGACAAT 5066
Db 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
Qy 5067 GGTCCAGCTTTTGTGGCCAGGTAAGTCAGGAGCTGCCCAAGATATTGGGATTGATTGG 5126
Db 921 GlyProAlaPheValAlaHisValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTyr 940
Qy 5127 AAATGCAATTGTGCATACAGACCCCAAGCTCAGGACAGGTAGAGAGATGAATAGAAC 5186
Db 941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
Qy 5187 ATTAAAGAGACCTTACTAAATTGACCGCGAGACTGCGGTAAATGATTGGATAGCTCTC 5246
Db 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
Qy 5247 CTGCCCTTGTGCTTTTGGGTAGGAACACCCCTCGACAGTTTGGGCTGACCCCTAT 5306
Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
Qy 5307 GAATTACTCTACGGGGGACCCCTTGGTGAATAATTGCTTCTGTACATAGTGTGAC 5366
Db 1001 GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
Qy 5367 GTGCTGCTTCCAGCCCTTGTCTTAGGCTCAAGGCACCTTGAGTGGGTGAGACAACGA 5426
Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
Qy 5427 GCGTGAAGCAACTCCCGGAGGCTACTCAGGAGAGGAGACTTGCAGATCCCACTCGT 5486
Db 1041 AlaTrpArgGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnIleProHisArg 1060
Qy 5487 TTCCAAGTGGAGATTCACTTACGTTAGACGCCACCGTGCAGGAAACCTCGAGACTCGG 5546
Db 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
Qy 5547 TGGNAGGCGCTTATCTCTACTTTTGACACCAACCGCTGTGAAGTCGAGGAATC 5606
Db 1081 TrpLysGlyProTyrLeuValLeuLeuThrProThrAlaValLysValGluGlyIle 1100
Qy 5607 TCACCTTGGATCCATGCATCCACGTTAAACCGGCGCACCTCCCGATTGGGGTGGAAA 5666
Db 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTyrLys 1120
Qy 5667 GCCAAAAGACTGAAAATCCCTTAAGCTTCGCTTCCATCGCTGGTTCCTTACTCTGTC 5726
Db 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
Qy 5727 AATAACCTCTCAGAC 5741
Db 1141 AsnAsnLeuSerAsp 1145
RESULT 10
Q4VFZ2_9GNMR PRELIMINARY; PERT; 1145 AA.
AC Q4VFZ2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Pol protein.
GN Name=pol;
OS Porcine endogenous retrovirus C/A.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gamma-retrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=286072;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C/A 4;
RA Denner J., Karlas A., Votteler Jr.;
RT "Nucleotide sequence and extended characterization of a high passage
RT human cell-adapted recombinant PERV-C/A.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY953542; AAY28928.1; -, Genomic DNA.
SQ SEQUENCE 1145 AA; 127925 MW; 62014B9629B04051 CRC64;

Alignment Scores:

Pred. No.:	0	Length:	1145
Score:	5986.00	Matches:	1134
Percent Similarity:	99.6%	Conservative:	5
Best Local Similarity:	99.1%	Mismatches:	5
Query Match:	40.9%	Indels:	0
DB:	2	Gaps:	0

US-10-723-552-3 (1-8132) x Q4VP22_9GNMR (1-1145)

Qy	2307	ATGGGTGCGCACAGGGCAACACAGATATCCATGCGACTACCCGAAACAGATTGACTTGGGA	2366
Db	1	MetGlyAlaThrGlyGlnGlnGlnTyrProTrpThrThrArgArgThrValAspLeuGly	20
Qy	2367	GTGGGACGGGTAAACCCACTCTCTTCTGCTCATCTAGTGCCCGACAGCCCTCTTAGGT	2426
Db	21	ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly	40
Qy	2427	AGAGACTTATTGACCAAGATCGGAGCACAAATTTCTTTGAAACAGGGAAACACAGAAGTG	2486
Db	41	ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal	60
Qy	2487	TCTGCAATAACAAACCTTATCAGTGTGTGACCTCCCAATPAGATGACGAATATCGACTA	2546
Db	61	SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu	80
Qy	2547	TACTCTCCCTAGTAAGCCCTGATCAAAATATACAAATCTGGTTGGGAACAGTTTCCCAA	2606
Db	81	TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln	100
Qy	2607	GCCTGGGCAGAAACCGCAGGATGGGTTTGGCAAGCAAGTTCCCCACCAAGTTATTCAA	2666
Db	101	AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln	120
Qy	2667	CTGAAGCCAGTGCCACACAGTGTCTCAGTACAGCAGTACCCCTTGAGTAAGAAGCTCAA	2726
Db	121	LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln	140
Qy	2727	GAAGGAATTCGGCCGCGATGCCAAGATTATCAACAGGCGATCTCTAGTTCCTGTCCAA	2786
Db	141	GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln	160
Qy	2787	TCTCCCTGGGAATCTCCCTGCTACCGGTTAGAAAGCTGGAGCTAATGACTATCGACCA	2846
Db	161	SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro	180
Qy	2847	GTACAGGACTTGAGAGAGGTCAATAAACGGGTGCGAGGATATACACCCCAACAGTCCCGAAC	2906
Db	181	ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn	200
Qy	2907	CCTTATAACCTCTTGTTGTCTCCACCCCAACGAGCTGGTATACAGTATTGGACTTA	2966
Db	201	ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTyrThrValLeuAspLeu	220
Qy	2967	AAGGATCGCTTCTTCTGCTGAGATTACCCCACTAGCAACCACTTTTGGCTTCGAA	3026
Db	221	LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu	240
Qy	3027	TGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGAGCCCGACTGCCCAAGGG	3086
Db	241	TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly	260
Qy	3087	TTCAAGAACTCCCGGACATCTTTTGACGAAGCCCTACACAGACCTGGCCAACTTCAGG	3146
Db	261	PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg	280
Qy	3147	ATCCACACCTCAGGTGACCTCTCCAGTACGTGATGACCTGCTTCTGGCGGGAGCC	3206
Db	281	IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla	300

Qy	3207	ACCAACAGAGCTGCTTAGAAGGCACGAAAGCAGCTACTGCTGGAAATGTCTGACCTAGGC	3266
Db	301	ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly	320
Qy	3267	TACAGAGCTCTCTTAAGAAGGCCAGATTGCGAGGAGAGGTAAACATACTTGGGTAC	3326
Db	321	TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr	340
Qy	3327	AGTTTGGGACGCGGACGCGATGGCTGACGAGGACGCGAAGAAACTGTAGTCCAGATA	3386
Db	341	SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle	360
Qy	3387	CCGGCCCAACACAGACCAACAAATGAGAGAGTTTTTGGGGACAGCTGGATTTTGCAGA	3446
Db	361	ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg	380
Qy	3447	CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCCACTCTACCCGCTAACCAAGAAAAA	3506
Db	381	LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLys	400
Qy	3507	GGGGAATTCTCTGGGCTCTGAGCACGAGAGCATTTGATGCTATCAAAAAGCCCTG	3566
Db	401	GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu	420
Qy	3567	CTGAGCGACCTGCTCTGGCCCTCCCTGACGTAACTAAACCCCTTACCCTTTATGTGAT	3626
Db	421	LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp	440
Qy	3627	GAGGTAAAGGAGTNGCCGGGAGTTTTTAACCCAAACCCCTAGGACCATGGAGAGACCT	3686
Db	441	GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro	460
Qy	3687	GTGCGCTACCTGTCAAAGAAGCTCGATCTGTAGCCAGTGGTGGCCCATATGCTCAAG	3746
Db	461	ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys	480
Qy	3747	GCTATCGCAGCTGTGGCCATATCTGGTCAAGGACGCTGACAAATGACTTTGGGACAGAAT	3806
Db	481	AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn	500
Qy	3807	ATACTGTATAGCCCCCGATGCTATGGAGAACATCGTTGGCAGCCCCCAGACCGATGG	3866
Db	501	IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp	520
Qy	3867	ATGACCAACCCCGCATGACCCACTCAAAAGCTGCTCTCACAGAGAGGCTCACGCTTC	3926
Db	521	MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe	540
Qy	3927	GCTCCACGACCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGAACCAATG	3986
Db	541	AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal	560
Qy	3987	ACTCATGATTGCCATCACTATTGATTGAGAGACTGGGGTCCGCAAGGACCTTACAGAC	4046
Db	561	ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp	580
Qy	4047	ATACCGCTGACTCGAGAAGTCTAACTGCTGCTCAACGCGAAGACAGCTATGTGTGCGAA	4106
Db	581	IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu	600
Qy	4107	GGTAAGAGGATGCTGGGGCGGGTGGTGGAGCGGACCCGCGACGATCTGGGCCAGCAGC	4166
Db	601	GlyLysArgMetAlaGlyAlaAlaValAspGlyThrArgThrIleTrpAlaSerSer	620
Qy	4167	CTCCCGAAGGAATTCAGCACAAAAGCTGAGCTCATGGCCCTCAGCGAAGCTTTCGGG	4226
Db	621	LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg	640
Qy	4227	CTGCGCGAAGGGAATTCATAAATTTATACGACAGCAGAGGTATGCTTTTGGACTGCA	4286
Db	641	LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla	660
Qy	4287	CACGTACATGGGGCCATCTATAAACAAGGGGGTGTGTTACCTCAGCAGGAGGGAATA	4346

```
Db 661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 4347 AAGAACAAGAGGAAATTTCTAAGCCTATTAGACCGCTACATTTACCAAAAAGGCTAGCT 4406
Db 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 700
QY 4407 ATTATACACTCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACCATG 4466
Db 701 IleIleHisCysProGlyHisGlnLysAlaLysAspProIleSerArgGlyAsnGlnMet 720
QY 4467 GCTGACCGGGTTGCCAAGCAGCGCCAGCGGTGTTAACTCTTCTGCTTATATAGAAATG 4526
Db 721 AlaAspArgValAlaLysGlnAlaGlnGlyValAsnLeuLeuProMetIleGluThr 740
QY 4527 CCCAAAGCCCCAGAACCCAGACGACGATACACCTTAGAAGACTGCGCAGAGATAAAAAG 4586
Db 741 ProLysAlaProGluProGlyArgGlnTyrThrLeuGluAspTyrGlnGluIleLysLys 760
QY 4587 ATAGACCAAGTTCTCTGAGACTCCGGAAGGACCTGCTATATCCTCAGATGGGAAGCAATC 4646
Db 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
QY 4647 CTGCCCCACAAAGAGGGTTAGAATATGTCCAACAGATACATCGTCTAACCCACTAGGA 4706
Db 781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 4707 ACTAAACACCTGCGAGCAGTGGTCGAGAACATCCCTTATCATGTTCTGAGGCTACACGGA 4766
Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
QY 4767 GTGCTGACTCGGTGGTCAACATGTGTGCGCTCCGACGCTGGTGTATGCTAAATCCTTCC 4826
Db 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY 4827 AGAATGCTCTCCAGGAAGAGACTAAGGGGAAGCACCAGCGCTCACTGGGAAGTGGAC 4886
Db 841 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTyrGluValAsp 860
QY 4887 TTCACCTAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTTTGTAGACACC 4946
Db 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
QY 4947 TTTTCAGATGGGTAGAGCTTATCCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAA 5006
Db 881 PheSerGlyTyrValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
QY 5007 AAAATACTGGAAGAAATTTTCCAAAGATTGGAATACCTTAAGGTAATAGGCTCAGACAAT 5066
Db 901 LysIleLeuGluGluIlePheProArgPheGlyLeuProLysValIleGlySerAspAsn 920
QY 5067 GGTCCAGCTTTTGTGCCAGGTAAAGTCAGGACTGGCCAGATATTTGGGATTTGATTGG 5126
Db 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTyr 940
QY 5127 AAATCGCATCTGCATACAGACCCCAAGCTCAGCAGAGTAGAGGATCAATAGAAC 5186
Db 941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY 5187 ATTAAGAGACCCCTTACTAAATTTGACCCGAGACTGGCGTGTAAATGATTGATAGCTCTC 5246
Db 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTyrIleAlaLeu 980
QY 5247 CTGCCCTTTGTGCTTTTATAGGTTAGGAACACCCCTGGACAGTTTGGCTGACCCCTTAT 5306
Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
QY 5307 GAATTAATCTACGGGGGACCCCCCACTGGTGTAGAAATGCTTCTGTACATAGTGTCTAC 5366
Db 1001 GluLeuLeuTyrGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
QY 5367 GTGCTGCTTTCCAGCCCTTTGTTCTCTAGGCTCAAGGCACCTTGAGTGGTGGTCAACGA 5426
```

```
Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTyrValArgGlnArg 1040
QY 5427 GCGTGGAGGCNACTCCGGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCACATCGT 5486
Db 1041 AlaTyrArgGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnIleProHisArg 1060
QY 5487 TTCCAAGTGGGAGATTTCAGTCTAGCTTAGACCGCCACCGTGCAGGAACCTCGAGACTCGG 5546
Db 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
QY 5547 TGGAAAGGCGCTTATCTCTGACTTTTGACCAACCAACGGCTGTGAAAGTCGAAGGAATC 5606
Db 1081 TrpLysGlyProTyrHisValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
QY 5607 TCCACCTGGATCCATGATCCACGTTAAACCGCGGCCACCTCCGATTCGGGTGGAAA 5666
Db 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTyrLys 1120
QY 5667 GCGAAAAGACTGAAAATCCCTTAAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTC 5726
Db 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
QY 5727 AATAACCTCTCA 5738
Db 1141 AsnAsnPheSer 1144

RESULT 11
Q8UMP5_9GAMR
ID Q8UMP5_9GAMR PRELIMINARY; PRT; 1146 AA.
AC Q8UMPS;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol.
OS Porcine endogenous retrovirus.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21568379; PubMed=11711616;
DOI=10.1128/JVI.75.24.12252-12265.2001;
RA Herring C., Quinn G., Bower R., Parsons N., Logan N.A., Brawley A.,
Elsome K., Whittam A., Fernandez-Suarez X.M., Cunningham D.,
Onions D., Langford G., Scobie L.;
RT "Mapping full-length porcine endogenous retroviruses in a large white
pig."
J. Virol. 75:12252-12265(2001).
RL EMBL; AY056035; AAL38193.1; -, Genomic_DNA.
DR HSP; P03355; 116J.
DR MEROPS; A02.020; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00075; KnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00879; RNASE_H; 1.
SQ SEQUENCE 1146 AA; 128417 MW; 1244B29F5C270DE CRC64;

Alignment Scores: 0 Length: 1146
Pred. No.: 1
```

Score:	5863.00	Matches:	1109
Percent Similarity:	98.6%	Conservative:	17
Best Local Similarity:	97.1%	Mismatches:	16
Query Match:	40.1%	Indels:	0
DB:	2	Gaps:	0
US-10-723-552-3 (1-8132) x Q8UMP5_9GNR (1-1146)			
QY	2307	ATGGGTGCCAGCGGCAACACAGTATCCATGGACTACCGAAGACAGATTGACTTGGGA	2366
DB	1	MetGlyAlaThrGlyGlnArgGlnTyr:ProTrpThrThrArgArgThrValAspLeuGly	20
QY	2367	GTGGGACGGGTAAACCCACTGTTCTGGTCATACCTGAGTCCCGACACCCCTCTTAGGT	2426
DB	21	ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly	40
QY	2427	AGAGACTATTACCAAGATGGGAGCACAAATTTCTTTGAACAAGGGAACACAGAAGTG	2486
DB	41	ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal	60
QY	2487	TCTGCAAAATACAAACCTATCAGTGTGTGACCCCTCCAATTAGATGACGAATATCGACTA	2546
DB	61	SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu	80
QY	2547	TACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATCTGGTTGGAAAGTTCGCCAA	2606
DB	81	TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln	100
QY	2607	GCCTGGGCAAAACCGCAGGATGGGTTTGGCAAGCAAGTTCCTCCACAAAGTTATTCAA	2666
DB	101	AlaTrpAlaGluThrAlaGlyIleGlyLeuAlaLysGlnValProProGlnValIleGln	120
QY	2667	CTGAAGGCCAGTGCACACCAAGTGTCACTGAGTACAGTACCTGAGTAAAGAGCTCAA	2726
DB	121	LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg	140
QY	2727	GAGGAATTCGGCGCATGTCGCAAGATTATCAACACAGGCGATCTCTGCTGTCCAA	2786
DB	141	GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuAlaProValGln	160
QY	2787	TCTCCCTGGAACTCCCTGCTACCGGTTAGAAAGCTGGGACTAATGACTATCGACCA	2846
DB	161	SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro	180
QY	2847	GTACAGGACTGAGAGAGTCAATAACGGGTGAGGATATACACCAACAGTCCCGCAAC	2906
DB	181	ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn	200
QY	2907	CCTTATAACTCTTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGACTTA	2966
DB	201	ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeu	220
QY	2967	AAGGATGCCCTTCTGCTGAGATTACACCCCACTAGCCAACTTTTGGCTTCGAA	3026
DB	221	LysAspAlaPheCysLeuArgGluHisProThrSerGlnProLeuPheAlaPheGlu	240
QY	3027	TGAGAGATCCAGGTACGGGAAGAACCGGGCAGTCTACCTGGACCCGACTGCCCCAAGGG	3086
DB	241	TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly	260
QY	3087	TTCAGAAGCTCCCGACCACTTTTGACGAAGCCCTACACAGACCTGGCCCACTTCAGG	3146
DB	261	PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg	280
QY	3147	ATCCAACACCTCAGGTGACCTCTCCAGTAGTGATGATGACCTGCTTCTGGGGGAGCC	3206
DB	281	IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla	300
QY	3207	ACCAACAGAGCTGCTTAGAAGGCACGAAGGCACTACTGCTGGAATTGCTGACCTAGGC	3266
DB	301	ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly	320
QY	3267	TACAGAGCCTCTGCTAAGAGGCCAGATTTGCGAGGAGAGGTAAACATACTTTGGGGTAC	3326
DB	321	TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr	340
QY	3327	AGTTTGGGACCGGCGAGCGATGCTCAGCGAGCGACGGAGAAACCTAGTCCAGATA	3386
DB	341	SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle	360
QY	3387	CCGGCCCAACACAGCAACAAATGAGAGAGTGTGTTGGGGACAGCTGGATTTTGACA	3446
DB	361	ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg	380
QY	3447	CTGTGGATCCCGGGTTTGGCAGCTTAGCAGCCCCACTCTACCCGCTAAACCAAGAAAA	3506
DB	381	LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys	400
QY	3507	GGGGAATTCCTCGGGCTCTGAGCAGCACCAAGAGCATTTGATGCTATCAAAAGGCCCTG	3566
DB	401	GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu	420
QY	3567	CTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTACCTTTTATGTGAT	3626
DB	421	LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp	440
QY	3627	GAGGTAAGGGAGTAGCCCGGGAGTGTAAACCCAAACCTTAGGACCATGGAGAAGACCT	3686
DB	441	GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro	460
QY	3687	GTCCCTACCTGTCAAAGAGCTCGATCCCTGATGCGAGTGTGGGCCCATATGCTCAAG	3746
DB	461	ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys	480
QY	3747	GCTATCGCAGCTGTGGCCATCTGCTCAAGAGCGCTGACAAATTTGACTTTGGGACAGAAT	3806
DB	481	AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuArgGlnAsn	500
QY	3807	ATAACTGTATAGCCCCCATGCAATTGGAGAAACATCGTTCCGCGAGCCCCAGACCGATGG	3866
DB	501	IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp	520
QY	3867	ATGACCAAGCCCGCATGACCCACTCATCAAGCTGCTCTTCACAGAGAGGCTCACCTTC	3926
DB	521	MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe	540
QY	3927	GCTCCACCAAGCCCTCTCAACCTGCGACCTCTCTGCTCAAGAGACTGATGAACAGTG	3986
DB	541	AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluIleThrAspGluProVal	560
QY	3987	ACTCATGATTGCCATCAACTATTGATTGAGAGACTGGGGTCCGCAAGGACCTTACAGAC	4046
DB	561	ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp	580
QY	4047	ATACCGCTGACTGGAGAAGTCTTAACCTGGTTCACCTGCGGAGAGAGCTATGTTGGTGA	4106
DB	581	IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerTyrValValGlu	600
QY	4107	GGTAAGAGATGCTGGGCGCGGTGGTGGAGCGGACCCGACACCATCTGGGCCAGCAGC	4166
DB	601	GlyLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAlaSerSer	620
QY	4167	CTGCCGGAAGGAACCTTACGACCAAAAGCTGAGCTCATGGCCCTCACGCAAGCTTTCGG	4226
DB	621	LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg	640
QY	4227	CTGGCCGAAGGGAATCCATAACATTTATACGAGCAGCGAGGTATGCTTTGGGACTGCA	4286
DB	641	LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla	660
QY	4287	CAGTACATGGGCCCCATCTATAACAAAGGGGTTCTGCTTACCTCAGCAGGAGGGAAATA	4346
DB	661	HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle	680
QY	4347	AAGAACAAAGAGGAAATTTCTAAGCCCTATTAGAAGCCGTACATTTACCAAAAGGCTAGCT	4406

681	Db	LyAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla	700
4407	Qy	ATTATACACTGTCTCTGCACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATG	4466
701	Db		720
4467	Qy	IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet	4526
721	Db		740
4527	Qy	GCTGACCGGGTTGCCAAGCAGGAGGCCCGCCAGGGTGTAACTTCTGCCTATATATAGAAATG	4586
741	Db	AlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThr	4646
4587	Qy		4666
761	Db	ATAGACAGATTCTCTGAGACTCCGGAAGGACCTGTCTATACCTCAGATGGGAAGGAAATC	780
4647	Qy	CTGCCCCAACAAAGAAGGGTTAGATATGTCCAAACAGATACATCGTCTAACCCACCTTAGGA	4700
781	Db	LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly	800
4707	Qy	ACTAAACACCTGCAGCAGTAGTGGTCAGAACATCCCTTATCATCTTCTGAGGCTACACAGA	4766
801	Db	ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly	820
4767	Qy	GTGGCTCACTCGGTGGTCAAAACATTGTGCCCTGCCAGCTGGTTAATGCTAATCTCTTC	4826
821	Db	ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer	840
4827	Qy	AGATGGCTCCAGGGAAGAGACTAAGGGGAAGCCACCAGCGCTCACTCGGGAAGTGGAC	4886
841	Db	ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisnTrpGluValAsp	860
4887	Qy	TTCTCAGAGTAAAGCCGGCTAAATACCGGAAAACAATACCTATTGGTTTGTGTAGACACC	4946
861	Db	PheThrGluValLysProAlaLysTyTrGlyAsnLysTyTrLeuLeuValPheValAspThr	880
4947	Qy	TTTTTCAGATGGGTAGAGGCTTATCCCTPACTAAGAAAGAGAGACTTCAACCGTGGTGGCTAAA	5006
881	Db	PheSerGlyTrpValGluAlaTyProThrLysLysGluThrSerThrValValAlaLys	900
5007	Qy	AAAATACTGGAAGAAATTTTTTCCAAGATTTTGGAAATACCTAAGGTAAATPAGGGTCAGACAAT	5066
901	Db	LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn	920
5067	Qy	GGTCCACCTTTTGTGTGCCAGTAAAGTCAGGAGCTGGCCCAAGATATTGGGGATTGATTGG	5126
921	Db	GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp	940
5127	Qy	AAACTGATTTGTGCATACAGACCCCAAGCTCAGGACAGGTAGAGAGATGAATAGAAC	5186
941	Db	LysLeuHisCysAlaTyraArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr	960
5187	Qy	ATTTAAAGAGACCCCTTACTAAATTGACCGCGGAGACTGGCGCTTAATATTGGATAGCTCTC	5246
961	Db	IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu	980
5247	Qy	CTGCCCTTTGTGCTTTTGTAGGGTTAGGAACACCCCTCGAACAGTTTGGGCTGACCCCTCAT	5306
981	Db	LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr	1000
5307	Qy	GAATTACTCTACGGGGGACCCCCCAATTGGTAGAAATCTCTCTGTACATAGTCTGTGAC	5366
1001	Db	GlnLeuLeuTyTrGlyGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp	1020
5367	Qy	GTGCTGCTTTCCACGACCTTTGTTCTCTAGGCTCAAGGCATCTTCAAGTGGGTGAGACAACGA	5426
1021	Db	ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg	1040
5427	Qy	GCGTGGAGGCAACTCCGGGAGGCCCTACTCAGGAGGAGAGACTTGCAGATCCCAATCGT	5486
1041	Db	AlaTrpLysGlnLeuArgGluAlaTyTrSerGlyGluGlyAspLeuGlnValProHisArg	1060

Qy	5487	TTCCAAGTGGGAGATTACGTCCTAGCTTAGACGCCACGGTCACGAAACCTCGAGACTCGG	554
Db	1061	PheGlnValGlyAspSerValTyrValArgHisArgAlaGlyAsnLeuGluThrArg	1080
Qy	5547	TGGAAGGGCCCTTATCTCGTACTTTTTCACACACACCAACGGCTGTGAAAGTCGAAGGAATC	5606
Db	1081	TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle	1100
Qy	5607	TCCACCTGGATCCATGCATCCCCAGTCTAAACCGCGCCACCTCCCGATTCCGGGTGGAAA	5666
Db	1101	SerThrTrpIleHisAlaSerHisValLysLeuAlaProProAspSerGlyTrpArg	1120
Qy	5667	GCGGAAAGACTGAAATCCCTTAAGCTTCGGCTCCATCGCTGGTTCCTTACTCTCTGTC	5726
Db	1121	AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn	1140
Qy	5727	AATAAC 5732	
Db	1141	AsnAsn 1142	
RESULT 12			
Q8UM99_9GAMR			
ID	Q8UM99_9GAMR	PRELIMINARY; PRT; 1144 AA.	
AC	Q8UM99_9		
DT	01-MAR-2002	(TREMBlrel. 20, Created)	
DT	01-MAR-2002	(TREMBlrel. 20, Last sequence update)	
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)	
DE	Pol protein.		
GN	Name=pol;		
OS	Porcine endogenous retrovirus.		
OC	Viruses; Retrofroid viruses; Retroviridae; Gammaretrovirus;		
OC	1-Mammalian type C virus group.		
OX	NCBI_TaxID=61673;		
ON	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=20219394; PubMed=10756014;		
EX	DOI=10.1128/JVI.74.9.4028-4038.2000;		
RA	Clauderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;		
RT	"Stabilishment and Characterization of molecular clones of porcine		
RT	endogenous retroviruses replicating on human cells.";		
RL	J. Virol. 74:4028-4038(2000).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
EX	MEDLINE=21851083; PubMed=11861838;		
EX	DOI=10.1128/JVI.76.6.2714-2720.2002;		
RA	Niebert M., Rogel-Gaillard C., Chardon P., Toenjes R.R.;		
RT	"Characterization of Chromosomally Assigned Replication-Competent Gamma		
RT	Porcine Endogenous Retroviruses Derived from a Large White Pig and		
RT	Expression in Human Cells.";		
RL	J Virol. 76:2714-2720(2002).		
DR	EMBL; AJ279056; CAC82502.1; -; Genomic_DNA.		
DR	HSSP; P03355; I16J.		
DR	MEROPS; A02.020; -.		
DR	GO; GO:0003677; F:DNA binding; IEA.		
DR	GO; GO:0004523; F:ribonuclease H activity; IEA.		
DR	GO; GO:0003723; F:RNA binding; IEA.		
DR	GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.		
DR	GO; GO:0006310; P:DNA recombination; IEA.		
DR	GO; GO:0006278; P:RNA-dependent DNA replication; IEA.		
DR	InterPro; IPR002156; RNaseH.		
DR	InterPro; IPR001584; Rve.		
DR	InterPro; IPR00477; RVTse.		
DR	Pfam; PF000075; RNaseH; 1.		
DR	Pfam; PF00665; rve; 1.		
DR	Pfam; PF000078; RVT.1; 1.		
DR	PROSITE; PS50879; RNASE H; 1.		
SQ	SEQUENCE 1144 AA; 127926 MW; 5F3AA5A5EFPB98038 CRC64;		
Alignment Scores:			
Pred. No.:	0	Length:	1144
Score:	5840.50	Matches:	1107
Percent Similarity:	97.8%	Conservative:	12

Best Local Similarity: 96.8% Mismatches: 24
Query Match: 39.9% Indels: 1
DB: 2 Gaps: 1

US-10-723-552-3 (1-8132) x Q8UM99_9GAMR (1-1144)

```
QY 2307 ATGGGTGCCACAGGGCAACAACAGTATCCATGGACTACCCGAAAGAACAGATTGACTGGGA 2366
DB 1 MetGlyAlaThrGlyGlnGlnTyrProTrpThrThrArgThrValAspLeuGly 20
QY 2367 GTGGGACGGGTAAACCCACTGTTCTGTCTATACCTGAGTCCCGCAGCACCCCTCTTAGGT 2426
DB 21 ValGlyArgValThrGlnSerPheLeuGlyIleProGluCysProGluProLeuLeuGly 40
QY 2427 AGAGACTTATTACCAAGATGGGAGCACAAATTTCTTTGAACAAGGGAACACAGAAGTG 2486
DB 41 ArgAspLeuLeuThrThrMetGlyArgGlnIlePheGluGlnGlyThrProGluVal 60
QY 2487 TCTGCAATAACAACCTATCACTGTGTGTGACCTCCAAATTAGATGACGAATATCGACTA 2546
DB 61 SerAlaAsnAsnLysProIleThrValIleThrLeuGlnLeuAspAspGluTyrArgLeu 80
QY 2547 TACTCTCCCTAGTAAAGCCGTGATCAAAATATACAAATTTCTGGTTGGAAACAGTTTCCCAA 2606
DB 81 TyrSerSerLeuValLysProAspGlnAsnIleGlnPheTrpLeuGlnGlnPheProGln 100
QY 2607 GCCTGGGCAGAAACCGCAGGAGTGGGTTTGGCAAGCAAGTTCCCCACCAAGTTATTCAA 2666
DB 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120
QY 2667 CTGNAGCCAGTGCACACAGTGTAGTCAGACAGTACCCCTTGAGTAAGAAGCTCAA 2726
DB 121 LeuLysAlaSerAlaAlaProValSerValArgGlnTyrProLeuSerLysGluAlaArg 140
QY 2727 GAAGGAATTCGGCGCATGTCCTCCCAAGAGTAAATCAACAGGSCATCTAGTTCCTGTCCAA 2786
DB 141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnIleLeuValProValGln 160
QY 2787 TCTCCCTGGGAATACTCCCTGCTACCGGTAGAAAGCCTGGGACTAATGACTATCGACCA 2846
DB 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
QY 2847 GTACAGACTTGAGAGAGTCAATAAAGGGTGCAGGATATACACCAACAGTCCCGAAC 2906
DB 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200
QY 2907 CCTTATAACCTCTGTGTGCTCTCCCAACCGAGCTGGTATACAGTATTGGACTTA 2966
DB 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu 220
QY 2967 AAGGATGCCCTTCTTCTGCTGAGATTACACCCCACTAGCAACACACTTTTGTCTCGAA 3026
DB 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
QY 3027 TGGAGAGATCCAGGTACGGGAAGAACCGGCAGCTCACTGGACCCGACTGCCCCCAAGGG 3086
DB 241 TrpArgAspProGlyAlaGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
QY 3087 TTCAAGAATCCCGACCATCTTTGAGAGCCCTTACAGAGACCTGGGCCCACTTCAGG 3146
DB 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280
QY 3147 ATCAACACACCTCAGGTGACCCCTCCCTCAGTACGTGATCACTGCTTCTGGCGGGAGCC 3206
DB 281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 300
QY 3207 ACCAAACAGACTGTCTTAGAAGCAGCAAGGCACTACTGCTGGAATTGCTGACCTAGGC 3266
DB 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 320
QY 3267 TACAGACCTCTCTTAAGAGGCCAGATTGACGAGAGAGTAACTACTTGGGGTAC 3326
DB 321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 340
```

```
QY 3327 AGTTTGGGACGGGACGATGCTGACGAGGACGGAAGAAAACCTGTAGTCCAGATA 3386
DB SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysArgThrValValGlnIle 360
QY 3387 CCGGCCCAACACAGCAACAAATGAGAGAGTTTTTGGGGAAGCTGTGGATTTTGCAGA 3446
DB ProAlaProThrThrAlaLysGlnValArgLysPheLeuGlyThrAlaGlyPheCysArg 380
QY 3447 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAACCAACCAAGAAAA 3506
DB LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 400
QY 3507 GGGGAATTCCTCGGGCTCTGAGCACCAGAGGCATTTGATGTATCAAAAAGGCCCTG 3566
DB 401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
QY 3567 CTGAGCGCACTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTTACCTTTATGTGAT 3626
DB LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440
QY 3627 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCCCTAGGACCATCGAGAGACCT 3686
DB GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 460
QY 3687 GTCGCCTTACCTGTCAAGAAAGCTCGATCCTGTAGCCAGTGTGGCCCATATGCTCAAG 3746
DB ValAlaTyrLeuSerLysLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys 480
QY 3747 GCTATGCGACGTGTGGCCATACCTGGTCAAGGACGCTGACAAATTTGACTTTGGGACAGA 3806
DB AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500
QY 3807 ATAACTGTAATAGCCCCCATGCTGGGAAACATCGTTCGGCAGCCCCCAGACCCGATGG 3866
DB IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 520
QY 3867 ATGACCAACCCCGCATGACCCACTACTCAAAAGCCTGCTTCTCAGAGAGAGGTCACGCTC 3926
DB MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540
QY 3927 GCTCCACAGCCCTCTCAACCCCTGCACTCTTCTGCTCAAGAGACTGTATGAACCAAGT 3986
DB AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 3987 ACTCATGATTGCCATCAACTATTGATTGAGGACACTGGGGTCCGCAAGGACCTTACAGAC 4046
DB ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580
QY 4047 ATACCCGTGACTGGGAAGTGCTAACTGTTCACTGACGGAAGCAGCTATGTGGTGAA 4106
DB IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValLys 600
QY 4107 GGTAAAGAGATGCTGGGGCGGGTGTGTGACCGGACCCGCAACGATCTCGGCCACGACG 4166
DB GlyLysArgMetAlaGlyProProValValAspGlyThrArgThrIleTrpAlaSerSer 620
QY 4167 CTGCGGGAAGGAACTTCAGCACAAAAGGCTGAGCTCATGCGCCCTCAGCAGCTTTGCGG 4226
DB LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 4227 CTGCGCGAAGGGAAATCCATAAATTTATACCGACAGCAGGTATGCTTTGGACGTGCA 4286
DB LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
QY 4287 CAGTACATGGGCCATCTATAAACAAAGGGGTGCTTACCTCAGCAGGAGGGAAATA 4346
DB HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 4347 AAGAAACAAGAGGAAATCTTAGCCCTATTAGAGCCGTACATTTTACCAAAAAGGCTAGCT 4406
DB LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 700
```

4407 ATTATACACTGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACACAGATG 4466
|||||
Db
701 IleIleHisCysProGlyHisGlnIleValAlaIleValSerProIleSerArgGlyValGlnMet 720
|||||
Qy
4467 GCTACCGGGTTGCCAGCAGCCAGCGGTGTTAACTCTCTGCTTATATAGAAATG 4526
|||||
Db
721 AlaAspArgValAlaIleValGlnAlaAlaGlnGlyValAenLeuLeuProMetIleGluThr 740
|||||
Qy
4527 CCCAAAGCCCGAGAACCCAGACGACGTACACCTCTAGAGACTGCCAGAGATATAAAG 4586
|||||
Db
741 ProIleValProGluProGlyArgGlnIleThrLeuGluAspTrpGlnGluIleValLys 760
|||||
Qy
4587 ATACACGAGTCTCTGAGACTCCGAGAGGACCTGCTATACCTCAGATGGGAAGAAATC 4646
|||||
Db
761 IleAspGlnPheSerGluThrProGluGlyThrCysThrThrSerAspGlyLysGluIle 780
|||||
Qy
4647 CTGCCCCCAAGAGGGTTAGAAATATGTCCAAACAGATACATCGTCTAAACCCACTAGGA 4706
|||||
Db
781 LeuProHisIleValGlyLeuGluThrValGlnGlnIleHisArgLeuLeuThrHisLeuGly 800
|||||
Qy
4707 ACTAAACACCTGCGACGAGTGGTTCAGAACATCCCTTATCATGTCTCAGGCTACACGGA 4766
|||||
Db
801 ThrLysHisLeuGlnGlnLeuValArgThrSerProIleHisValLeuArgLeuProGly 820
|||||
Qy
4767 GTGCTGACTCGGTGGTCAAACTTTGTGCGCCCTGCCAGCTGGTGTAAATGCTAACTCTCC 4826
|||||
Db
821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAenAlaAsnProSer 840
|||||
Qy
4827 AGAATGCTCTCAGGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCCTAGGGAAGTGGAC 4886
|||||
Db
841 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
|||||
Qy
4887 TTCACTGAGTAAAGCGGCTAAATACGGAACAAATACCTATTGGTTTTGTAGACACC 4946
|||||
Db
861 PheThrGluValLysProAlaLysThrGlyAsnLysThrLeuLeuValPheValAspThr 880
|||||
Qy
4947 TTTTCAGATGGGTAGAGGCTTATCTACTAAGAAAGAGACTCAACCGTGGTGGCTAAA 5006
|||||
Db
881 PheSerGlyTrpValGluAlaThrProThrLysLysGluThrSerThrValValAlaLys 900
|||||
Qy
5007 AAAATACTGGAAGAAATTTTCCAGATTTTCCAGATCTTAAGGTAATAGGTCACAGCAAT 5066
|||||
Db
901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
|||||
Qy
5067 GGTCCAGCTTTGTCGCCAGTAAGTCAGGACTGGCAGATATTTGGGATTTGATGG 5126
|||||
Db
921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
|||||
Qy
5127 AAATGCAATTGTGCATACAGACCCCAAGCTCAGACAGGTAGAGAGATGAATAGAAC 5186
|||||
Db
941 LysLeuHisCysAlaIleValArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
|||||
Qy
5187 ATTAAGAGACCCCTTACTAAATTCACCGCGAGACTGGCGTTAATGATTTGATGCTCTC 5246
|||||
Db
961 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAenAspTrpIleAlaLeu 980
|||||
Qy
5247 CTGCCCTTTGCTTTTAGGGTTAGGNACACCCCTGACAGATTTGGCTGACCCCTTAT 5306
|||||
Db
981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProIle 1000
|||||
Qy
5307 GAATTTACTCTACGGGGGACCCCCCACTTGTGTAGAAATTTGCTCTGTACATAGTGTGAC 5366
|||||
Db
1001 GluLeuLeuThrGlyGlyProProPheLeuAlaGluIleAlaPheAlaHisSerAlaAsp 1020
|||||
Qy
5367 GTGCTGCTTTCCAGCCCTTTGCTCTAGGCTCAAGGCACTTTGAGTGGGTGAGACAAGA 5426
|||||
Db
1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluThrValArgGlnArg 1040
|||||
Qy
5427 GCGTGGAGGCAATCCCGGAGCCCTACTCAGAGGAGAGACTTCGATCCCATCGT 5486
|||||
Db
1041 AlaTrpLysGlnLeuArgGluAlaThrSer---GlyGlyAspLeuGlnValProHisArg 1059
|||||
Qy
5487 TTCCAAGTGGGAGATTTCAGTCTAGCTAGACGCCACCGTCGAGGAAACCTCGAGACTCGG 5546
|||||

Db
1060 PheGlnValGlyAspSerValTyValArgArgHisArgAlaGlyAsnLeuGluThrArg 1079
|||||
Qy
5547 TGGAGGGCCCTTATCTCTGACTTTTGCACACACCAACGGCTGTGAAGTGCAGAGGAATC 5606
|||||
Db
1080 TrpLysGlyProIleLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1099
|||||
Qy
5607 TCCACCTGGATCCATGCCATCCAGTTAAACCGCGCCACCTCCCGATTCCGGGTGGAAA 5666
|||||
Db
1100 ProThrTrpIleHisAlaPheHisValLysProAlaProSerAspSerGlyTrpLys 1119
|||||
Qy
5667 GCCGAAAGACTCGAAATCCCTTAACTTTCGCTCCATCGCTGGTTCCTTACTCTGTC 5726
|||||
Db
1120 AlaGlnLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProIleSerVal 1139
|||||
Qy
5727 AATAACCTCTCA 5738
|||||
Db
1140 AsnAsnSerSer 1143
|||||
RESULT 13
Q8UM96_9GAMR
ID Q8UM96_9GAMR PRELIMINARY; PRT; 1147 AA.
AC Q8UM96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol protein.
GN Name:pol;
OS Porcine endogenous retrovirus.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20219394; PubMed=10756014;
DOI=10.1128/JVI.74.9.4028-4038.2000;
RA Cauderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;
RT "Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells."
RL J. Virol. 74:4028-4038 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21851083; PubMed=11861838;
DOI=10.1128/JVI.76.6.2714-2720.2002;
RA Niebert M., Rogel-Gaillard C., Chardon P., Toenjes R.R.;
RT "Characterization of Chromosomally Assigned Replication-Competent Gamma Porcine Endogenous Retroviruses Derived from a Large White Pig and Expression in Human Cells."
RL J. Virol. 76:2714-2720 (2002).
DR EMBL; AJ279057; CAC82505.2; -; Genomic_DNA.
DR MEROPS; A02.020; -;
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR00477; RVTse.
DR Pfam; PF000075; KnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50879; RNASE_H; 1.
SO SEQUENCE 1147 AA; 128404 MW; C6374B0CB460901D CRC64;
Alignment Scores:
Pred. No.: 0 Length: 1147

Score: 5831.00 Matches: 1104
Percent Similarity: 98.4% Conservative: 20
Best Local Similarity: 96.7% Mismatches: 18
Query Match: 39.8% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3 (1-8132) x Q8UM96_9GMR (1-1147)

```
QY 2307 ATGGGTGCCAGGCGCAACACAGTATCCATGAGTACCCGAGAACAGATTGACTTGGGA 2366
DB 1 MetGlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgArgThrValAspLeuVal 20
QY 2367 GTGGGAGGGTAACCCACTGTTCTGTGTATCTAGTGTGCCAGACCCCTCTTAGGT 2426
DB 21 ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly 40
QY 2427 AGAGACTTATTGACCAAGATGGGAGCACAAATTTCTTTGACCAAGGGAACCAAGAGTG 2486
DB 41 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 60
QY 2487 TCTGCAAAATACAAACCTTATCAGTGTGTGACCCCTCCAAATTAGATGACGAATATCGACTA 2546
DB 61 SerValAspLeuLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 80
QY 2547 TACTCTCCCTAGTAAGCTGTATCAAAATATACAAATCTGTTGGTAACAGTTTCCCAA 2606
DB 81 TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln 100
QY 2607 GCCTGGGACAAACCGGAGGATGGTTGGCAAGCAAGTTCCCCACCAAGTTATTCAA 2666
DB 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnValIleGln 120
QY 2667 CTGAAGGCCAGTGCACACAGTGTCTAGTCAGACAGTACCCCTTGAGTAAAGAGCTCAA 2726
DB 121 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg 140
QY 2727 GAAGGAATTCGGCGCAGTGTCCAAAGATTAATCAACACAGGCAATCTAGTTCCTGTCCAA 2786
DB 141 GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyLeuValProValGln 160
QY 2787 TCTCCCTGGAAATCTCCCTGCTACCGGTTAGAAAGCTGGGACTAATGACTATCGACCA 2846
DB 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
QY 2847 GTACAGACTTGAGAGAGTCAATAAACGGGTGCAGGATATACACCCAAACAGTCCCGAAC 2906
DB 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200
QY 2907 CCTTATAACCTTGTGTGTCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGACTTA 2966
DB 201 ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeu 220
QY 2967 AAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCAAACCACTTTTGGCTTCGAA 3026
DB 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
QY 3027 TGGAGAGATCCAGTACGGGAAGAACCGGCGAGCTCACCTGGACCCGCTGCCCAAGGG 3086
DB 241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
QY 3087 TTCAGAATCTCCCGACCATCTTTGACGAAGCCCTACACAGACCTGGCCAACTTCAGG 3146
DB 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280
QY 3147 ATCCAACACCTCAGGTGACCTCTCTCAGTACGTGATGACCTGCTTCTGGGGGAGGCC 3206
DB 281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 300
QY 3207 ACCAAACAGACTGTCTTAGAAGGACGAGGCACTACTGCTGGAATTTGCTGACCTAGGC 3266
DB 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuGlnLeuSerAspLeu 320
QY 3267 TACAGAGCCTCTCTGAAGAGGCCGAGATTGTGAGGAGAGAGGTAAACATCTTGGGGTAC 3326
```

```
DB 321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 340
QY 3327 AGTTTGGGAGCGGCGAGCGATGCTGACGAGGACGAGGAGAAACCTGTAGTCCAGATA 3386
DB 341 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 360
QY 3387 CCGGCCCAACACAGCAACAAATAGAGAGATTTTTTGGGACAGCTGGATTTTGCAGA 3446
DB 361 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 380
QY 3447 CTGTGGATCCGGGGTTTGGACCTTAGCAGCCCCACTCTACCCGCTAACCAAGAAAAA 3506
DB 381 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 400
QY 3507 GGGGAATTCCTCGGGCTCTGAGCACCAGAGGCATTTCATGCTATCAAAAGGCCCTG 3566
DB 401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
QY 3567 CTGAGCGCAGCTGCTCTGGCCCTCCCTGAGCTAACTAAACCTTTACCTTTATGTGGAT 3626
DB 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440
QY 3627 GAGCTTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGACCATGGAGAAGACCT 3686
DB 441 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 460
QY 3687 GTCCCTACTCTGTCAAGAAAGCTCGATCTCTGAGCAGTGTGGCCCATATGCTCGAAG 3746
DB 461 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLys 480
QY 3747 GCTATCGACCTGTGGCCATCTACTGTCAGAGCAGCTCAAAATGACTTTGGGACAGAA 3806
DB 481 AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500
QY 3807 ATAACCTGTAATAGCCCCCATGCAATTGGAGAACATCGTTCCGAGAGCCCGACAGCTGG 3866
DB 501 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 520
QY 3867 ATCACCACCGCCGATGACCCACTCAAGCCTGCTTCTCAGAGAGGGTCAACCTTC 3926
DB 521 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540
QY 3927 GCTCCACCGCCGCTCTCAACCCCTGCCACTCTCTGCTCAAGAGAGCTGATGAACAGTG 3986
DB 541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 3987 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGTCCGCAAGGACCTTTACAGAC 4046
DB 561 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580
QY 4047 ATACCGCTGACTGGAGAGTGCTAACTGTTTCACTGACCGAGAGCAGCTATGTTGGTGA 4106
DB 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
QY 4107 GGTAAAGAGGATGCTGGGGCGGCTGGTGGAGCGGACCCGACGATCTCGGCCAGCAGC 4166
DB 601 GlyLysArgMetAlaGlyAlaAlaValAspGlyThrArgThrIle**AlaSerSer 620
QY 4167 CTCCCGGAAGGAACCTTCAGCACAAAAGGCTGAGCTCATGGCCCTCAGCAGAGCTTTCGG 4226
DB 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 4227 CTGGCCGAGGGAATCCATAACATTTATACGAGCAGGAGGTATGCTTTGGGACTGCA 4286
DB 641 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
QY 4287 CAGTACATGGGGCCCATCTATAACAAAGGGGTGCTTACCTCAGCAGGAGGAGGAATA 4346
DB 661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 4347 AAGAACAAAGAGGAATTTCTAACCTTATTAGAGCCGTACATTTACCAAAAGGCTAGCT 4406
```

```
Db      LysAsnLysAspGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 700
Qy      4407 ATTATACACTGTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACCATG 4466
Db      701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
Qy      4467 GCTCAGCGGTGGCCAGCAGCGAGCCAGGGTGTAAACCTTCTGCTTATATAAGAAATG 4536
Db      721 AlaAspArgIleAlaLysGlnAlaGlnAlaValAsnLeuLeuProIleIleGluThr 740
Qy      4527 CCCAAAGCCCAGAACCCAGACGACGATACACCCCTAGAAAGCTGGCAAGATATAAAAG 4586
Db      741 ProLysAlaProGluProArgArgGlnThrLeuGluAspTrpGlnGluIleLysLys 760
Qy      4587 ATAGACAGTCTCTGAGACTCCGAGAGGACCTGCTATACCTCAGATCGGAAGGAATC 4646
Db      761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 780
Qy      4647 CTGCCCCACAAGAGGGTTAGAATATGTCACACAGATACATGCTCTAAACCCACTAGGA 4706
Db      781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
Qy      4707 ACTAAACACCTGCAGCAGTGTGGTCAGAACATCCCTTATCATGTCTCAGGCTACACGGA 4766
Db      801 ThrLysHisLeu***GlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
Qy      4767 GTGCTGACTCGGTGGTCAAAATGTGTGCCCTGCCAGCTGGTTAATGCTTAATCCTTCC 4826
Db      821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
Qy      4827 AGAATGCTCCAGGAGAGAGACTAAGGGAGCCACCCAGCGCTCACTGGGAAGTGGAC 4886
Db      841 LysValProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
Qy      4887 TTCCTAGGTAAAGCCGGCTAAATACAGAAACAAATACCTATTGGTTTTGTAGACACC 4946
Db      861 PheThrLysValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
Qy      4947 TTTTCAGATGGGTAGAGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 5006
Db      881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
Qy      5007 AAAATACTGGAAGAAATTTTCCAAGATTTCGAATACCTAAGGTAAATAGGCTCAGACAAT 5066
Db      901 LysIleLeuGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
Qy      5067 GGTCCAGCTTTGTGTCCCAAGTAAGTCAGGAGCTGCCCAAGATATTGGGATTGATTGG 5126
Db      921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
Qy      5127 AAATGCAATTGTGCATACAGACCCCAAGCTCAGGACGAGTAGAGAGATGAATAGAAC 5186
Db      941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
Qy      5187 ATTAAGAGACCCCTTACTAAATTCAGCCGCGAGACTGGCGTTAATGATTGATAGCTCTC 5246
Db      961 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu 980
Qy      5247 CTGCCCTTTGTGCTTTTGTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTCAGCCCTTAT 5306
Db      981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
Qy      5307 GAATTACTCTACGGGGAGACCCCCCAATTGTTAGAAATTTGTTCTGTATACATGCTGAC 5366
Db      1001 GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaProValHisSerAlaAsp 1020
Qy      5367 GTGCTGCTTTCCAGCCTTCTTCTCAGGCTCAAGGCACCTTGAGTGGTGGTCAGACAAGA 5426
Db      1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
Qy      5427 GCGTGGAGGCAACTCCGGGAGCGCTTACTCAGGAGGAGGAGACTTGCAGATCCCACTCGT 5486
Db      1041 AlaTrpLysGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnValProHisArg 1060
```

```
Qy      5487 TTCCAAGTGGAGATTTCAGTCTAGTTAGACGCCACCGTCGAGGAAACCTCGAGACTCGG 5546
Db      1061 PheGlnValGlyAspSerValTyrValArgHisArgAlaGlyAsnLeuGluThrArg 1080
Qy      5547 TGAAGGGCCCTTATCTCGTACTTTTACCACACACCAACGGCTGTGAAAGTCGAAGGAATC 5606
Db      1081 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
Qy      5607 TCACACTGGATTCATGCCATCCACGTTAAACCGGCGCCACCTCCCGATTTCGGGTGAAA 5666
Db      1101 SerThrTrpIleHisAlaSerHisValLysLeuAlaProProAspSerGlyTrpArg 1120
Qy      5667 GCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGCGTTCCTTACTCTCTGTC 5726
Db      1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn 1140
Qy      5727 AATAAC 5732
Db      1141 AsnAsn 1142

RESULT 14
Q8Q6U4_9GAMR
ID      Q8Q6U4_9GAMR PRELIMINARY; PRT; 1142 AA.
AC      Q8Q6U4;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Pol protein.
GN      Name=pol;
OS      Porcine endogenous retrovirus.
OC      Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC      1-Mammalian type C virus group.
OX      NCBI_TaxID=61673;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=21851083; PubMed=11861838;
RX      DOI=10.1128/JVI.76.6.2714-2720.2002;
RA      Niebert M., Rogel-Gaillard C., Chardon P., Toenjes R.R.;
RT      "Characterization of Chromosomally Assigned Replication-Competent Gamma
RT      Porcine Endogenous Retroviruses Derived from a Large White Pig and
RT      Expression in Human Cells.";
RL      J. Virol. 76:2714-2720(2002).
DR      EMBL; AF435967; AAL87856.1; -; Genomic_DNA.
DR      HSP; P03355; IMWL.
DR      MEROPS; A02.020; -.
DR      GO; GO:0003677; F:DNA binding; IEA.
DR      GO; GO:0004523; F:ribonuclease H activity; IEA.
DR      GO; GO:0003723; F:RNA binding; IEA.
DR      GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR      GO; GO:0006310; P:DNA recombination; IEA.
DR      GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR      InterPro; IPR001584; Rve.
DR      InterPro; IPR00477; RVTse.
DR      Pfam; PF00075; RnaseH; 1.
DR      Pfam; PF00665; rve; 1.
DR      Pfam; PF00078; RVT_1; 1.
DR      PROSITE; PS00879; RNASE H; 1.
SQ      SEQUENCE 1142 AA; 127712 MW; B0F55CA742AAC98F CRC64;

Alignment Scores:
Pred. No.:      6.19e-314      Length:      1142
Score:          5466.50      Matches:    1042
Percent Similarity: 93.4%      Conservative: 26
Best Local Similarity: 91.1%      Mismatches: 73
Query Match:    37.3%      Indels:     3
DB:             2              Gaps:       2

US-10-723-552-3 (1-8132) x Q8Q6U4_9GAMR (1-1142)
Qy      2307 ATGGGTGCCAGGCGCAACACAGTATCCATGGACTACCCGAGAACAGTTGACTTGGGA 2366
Db      |||||||
```

Db 1 MetGlyAlaThrGlyGlnGlnTyrProThrThrThrArgArgThrValAspLeuGly 20
Qy 2367 GTGGACGGTAACCACTCGTTCTGTGTCATACCTGAGTGCCACGACCCCTCTTAGGT 2426
Db 21 ValGlyArgValThrGlnSerPheLeuGlyIleProGluCysProGluProLeuLeuGly 40
Qy 2427 AGAGACTTATTGACCAAGATGGGACACAAATTTCTTTGAAACGAGGGAACCGAAGTG 2486
Db 41 ArgAspLeuLeuThrThrMetGlyArgGlnIlePhePheGluGlnGlyThrProGluVal 60
Qy 2487 TCTGCAATAACAAACCTTACCTGTTGTGACCTTCCAAATAGATGACGAATATCGACTA 2546
Db 61 SerAlaIaenAmLysProIleThrValLeuThrLeuGlnLeuAspGluTyrArgLeu 80
Qy 2547 TACTCTCCCTAGTAAGCCGTGATCAAAATATACAAATTTCTGGTGGAAACAGTTTCCCAA 2606
Db 81 PheSerProProValLysLeuAspGlnAsnIleGlnPheGlySerThrGlnPheProGln 100
Qy 2607 GCCTGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCTCCCAAGATTATTCAA 2666
Db 101 AlaLeuAlaGluProAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120
Qy 2667 CTGAAGGCCAGTGCACACCACTGTCTGTCAGACAGTACCTCTTGAGTAAAGAAAGCTCAA 2726
Db 121 LeuLysProSerLeuAlaProValProValArgGlnSerProPheSerLysGluAlaArg 140
Qy 2727 GAAGGAATTCGGCCGATGTCCAAAGATTATCAACAGGCGATCTTAGTTCTGTCCAA 2786
Db 141 GluGlyIleArgProHisIleValGlnArgLeuIleGlnGlnIleIleValProValGln 160
Qy 2787 TCTCCCTGGAAATCTCCCTGCTACCGGTTAGAAAGCTGGGACTAATGACTATCGACCA 2846
Db 161 SerProThrAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
Qy 2847 GTACAGACTTGACAGAGGTCAATAAAGGGTGCAGGATATACACCAACAGTCCCGAAC 2906
Db 181 ValGlnAspPheGluArgGlyGlnLysArgValGlnAspIleHisProThrValProAsn 200
Qy 2907 CTTTATAACCTCTTGTGTCTCTCCCAACCGGAGCTGGTATACAGATTGGACTTA 2966
Db 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTyrThrValLeuAspLeu 220
Qy 2967 AAGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCAAACCACTTTTGGCTTCGAA 3026
Db 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
Qy 3027 TGGAGAGATCCAGTACGGGAAGAACCGGCAGCTCACCTGGACCCGACTGCCCCAGGG 3086
Db 241 TrpArgAspProGlyAlaGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
Qy 3087 TTCAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGACCTGGCCAACTTCAGG 3146
Db 261 PheLysAsnPheProThrIlePheAspGlnAlaLeuHisArgAspLeuAlaAsnPheArg 280
Qy 3147 ATCAACACCTTAGGTGACCTCTCCAGTACGTGATGACCTGCTTCTGGCGGAGCC 3206
Db 281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuValLeuAlaAla 300
Qy 3207 ACCAAACAGACTCTTTAGAGGACCAAGGCATCTGCTGGAAATGCTGACCTAGGC 3266
Db 301 ThrLysGlnAspCysLeuGlnArgProLysGlyLeuLeuValGluLeuSerAspLeuGly 320
Qy 3267 TACAGACCTCTGTGAAGAGGCCAGATTTGACGAGAGAGGTAAATACTTTGGGGTAC 3326
Db 321 AspArgAlaPheGlyTyrLysAlaHisIleCysProThrGluValThrTyrLeuGlyTyr 340
Qy 3327 AGTTTGGGACGGGACGATGCTGACGAGGACGAGCAAGAACTAGTCCAGATA 3386
Db 341 ArgLeuArgGlyArgHisArgTyrLeuThrGluAlaProGlnThrThrValValGlnIle 360
Qy 3387 CCGGCCCAACCAACAGCAACAAATGACAGAGTTTGGGACAGCTGATTTGACAGA 3446
Db 361 ProGlyProThrProAlaLysGlnValArgGluPheLeuGlyThrValGlyPheCysArg 380

Qy 3447 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAA 3506
Db 381 LeuTrpIleProGlyPheAlaThrLeuProAlaProLeuTyrProLeuProLysGluLys 400
Qy 3507 GGGGAATTCCTCTGGGCTCTCGAGCACAGAGGCAATTTGATCTATCAAAAGCCCTG 3566
Db 401 GlyLupheSerTrpAlaLeuGlnHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
Qy 3567 CTGAGCCACCTGCTCTGGCCCTCCCTGAGCTACTAAACCTTTTACCTTTATGTGGAT 3626
Db 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLys-----ThrLeuTyrValAsp 438
Qy 3627 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGACCATGGAGAGACCT 3686
Db 439 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 458
Qy 3687 GTCGCTACTCTGCAAGAAAGCTCGATCTCTGTAGCCAGTGGTGGCCCATATGCTGAAG 3746
Db 459 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProLysCysLeuLys 478
Qy 3747 GCTATCCAGCTGTGGCCATCTGCTCAAGGAGCTGACAAATTTGACTTTGGGACAGAT 3806
Db 479 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 498
Qy 3807 ATAACGTGAATAGCCCCCATGCAATTGGAGAACATCTGTCGGCAGGCCCCAGACCGATGG 3866
Db 499 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 518
Qy 3867 ATGACCAACCCCGCATGACCCATATCAAGGCTGCTTCTCAGAGAGGGTCACTTC 3926
Db 519 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 538
Qy 3927 GCTCCACGACCGCTCTCAACCTGCGACCTCTGCTGCTCAAGAGAGCTGACCAAGTG 3986
Db 539 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProLeu 558
Qy 3987 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGTCGCGCAAGGACCTTACAGAC 4046
Db 559 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 578
Qy 4047 ATACCGCTGATCGGAGAGTGCTAACTGTTCTCACTGACGGAAGCACTATGTGGTGA 4106
Db 579 IleProLeuThrGlyGluProValThrTrpPheThrAspGlySerSerTyrLeuValGlu 598
Qy 4107 GGTAAAGAGATGCTGGCGCGGTGGTGCAGCGGACCCGACGATCTGGCCAGCAGC 4166
Db 599 GlyAsnLysMetAlaGlyAlaAlaValValAspArgThrProThrIleTrpGlyThrAsn 618
Qy 4167 CTCCCGAAGAACTTTCAGCAACAAAGGCTGAGCTCATGGCCCTCAGCGCAAGCTTTGCGG 4226
Db 619 LeuProGluArgThrSerSerGlnLysGlyGluLeuIleGlyLeuMetGlnAlaPheArg 638
Qy 4227 CTGGCGAAGGGAATTCATAAATTTATACGAGCAGCAGGTATGCTTTGCGACTGCA 4286
Db 639 LeuGlyGlnGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 658
Qy 4287 CAGTCATGCGGCCCATCTATAAACAAGGGGTGCTTACCTCAGCAGGAGGGAATA 4346
Db 659 HisValHisGlyAlaIleTyrThrGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 678
Qy 4347 AAGAAACAAAGAGAAAATTTTAAAGCCCTATTAGAACCGGTACATTTACCAAAAGCGTAGCT 4406
Db 679 LysAsnLysGluGluLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 698
Qy 4407 ATTATACATGTCCTCGACATCAAGAAAGCTAAAGATCTCATATCCAGAGAAACAGATG 4466
Db 699 IleIleHisCysProGlyHisGlnLysAlaLysAspProIleSerArgGlyAsnGlnMet 718
Qy 4467 GCTGACCGGGTTGCCAAGCAGGCGCCCGGGTGTAACTTCTGCTTAAATAGAAATG 4526
Db 719 AlaAspArgValAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProMetIleGluThr 738

```
Qy 4527 CCCAAGCCCCAGAACCCAGACGACAGTACACCCCTAGAGAGACTGCGAAGATATAAAAG 4586
Db 739 ProlysalProGluProGlyArgGlnThrLeuGluAspTrpGlnGluLeuLeuLeuLeu 758
Qy 4587 ATAGACCAAGTTCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATC 4646
Db 759 IleAspGlnPheSerGluThrProGluGlyThrCysfyrThrSerAspGlyLeuLeuLeu 778
Qy 4647 CTGCCCCCAAGAGGGTTAGATATGTCACAGATACATCGCTCAACCCACTAGGA 4706
Db 779 LeuProHisLeuGluGlyLeuGluThrValGlnGlnIleHisArgLeuThrHisLeuGly 798
Qy 4707 ACTAAACACCTGCAGCAGCTGCTCAGAACATCCCTTATCATGTTCTGAGGCTACACAGA 4766
Db 799 ThrLysHisLeuGlnLeuValArgThrSerProfyrHisValLeuArgLeuProGly 818
Qy 4767 GTGGCTGACTCGGTGGTCAAAATGTTGTGCCCTGCCACGCTGTTAATGCTTAATCTTCC 4826
Db 819 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 838
Qy 4827 AGAATGCTCCAGGAGAGACTAAGGGAGAGCCACCCAGGCGCTCACTGGGAAGTGCAC 4886
Db 839 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 858
Qy 4887 TTCACTGAGGTAAAGCCGCTAAATACGAAACAAATACCTATTGTTGTTGTAGACACC 4946
Db 859 PheThrGluValLysProAlaLysTrpGlyAsnLysfyrLeuLeuValPheValAspThr 878
Qy 4947 TTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 5006
Db 879 PheSerGlyTrpValGluAlaTrpProThrLysLysGlnThrSerThrValValAlaLys 898
Qy 5007 AAATACTCGAAGAAATTTTCCAGATTGCGAATTAAGTAAAGGTACAGCAAT 5066
Db 899 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 918
Qy 5067 GGTCAGCTTTGTTGCCAGGTAAAGTCAAGGACTGGCCAGATATTGGGATTGATTGG 5126
Db 919 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 938
Qy 5127 AAATGCTATTGTCATACAGACCCCAAGCTCAGGACAGGTAGAGAGATGAATAGAAC 5186
Db 939 LysLeuHisCysAlaTrpArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 958
Qy 5187 ATTAAGAGACCTTACTAAATTCACCGGAGACTGCGTTAATGATTGATAGCTCTC 5246
Db 959 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu 978
Qy 5247 CTGCCCCTTGTGCTTTTAGGGTTAGGAACACCCCTGGACAGTWTGGCTCACCCCTAT 5306
Db 979 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProfyr 998
Qy 5307 GAATTACTCTACGGGGGACCCCCCAATGTTGTTAGAAATTTGCTCTGTATAGTGTCTGAC 5366
Db 999 GluLeuLeuTyrglyGlyProProProLeuAlaGluIleAlaPheAlaHisSerAlaAsp 1018
Qy 5367 GTGCTGCTTCCAGCCTTTGCTTCTAGGCTCAAGGCACTTGTAGTGGTGGAGACAGA 5426
Db 1019 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1038
Qy 5427 GCGTGGAGCAACTCCGGGAGCGCTACTCAGGAGGAGAGACTTGCAGATCCCATCGT 5486
Db 1039 AlaTrpLysGlnLeuArgGluAlaTrpSer---GlyGlyAspLeuGlnValProHisArg 1057
Qy 5487 TTCCAAGTGGAGATTTCAGTCTACGTTAGACGCCCACTGCGAGGAACCTCGAGACTCG 5546
Db 1058 PheGlnValGlyAspSerValTrpValArgArgHisArgAlaGlyAsnLeuGluThrArg 1077
Qy 5547 TGGAAAGGCCCTTATCTGCTACTTTTACACCAACCAACCGCTGTCAAGTGCAGAGGATC 5606
Db 1078 TrpLysGlyProfyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1097
Qy 5607 TCCACCTGGATCCATCCATCCACGTTAAACCGGCGCCACCTCCCGATTCCGGGGTGGAA 5666
```

```
Db 1098 ProThrTrpIleHisAlaPheHisValLysProAlaProProSerAspSerGlyTrpLys 1117
Qy 5667 GCCGAAAAGACTGAAATCCCTTAAGCTTCGCCTCCATCGCGTGGTTCCTTACTCTGTC 5726
Db 1118 AlaGlnLysThrGluAsnProLeuLysLeuArgLeuHisArgValProfyrSerVal 1137
Qy 5727 AATAACCTCTCA 5738
Db 1138 AsnAsnSerSer 1141
RESULT 15
ID Q8Q6U7_9GAMR PRELIMINARY; PRT; 1139 AA.
AC Q8Q6U7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DB Pol protein.
GN Name: pol;
OS Porcine endogenous retrovirus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=21851083; PubMed=11861838;
RX DOI=10.1128/JVI.76.6.2714-2720.2002;
RA Niebert M., Rogel-Gaillard C., Chardon P., Toenjes R.R.;
RT "Characterization of Chromosomally Assigned Replication-Competent Gamma
RT Porcine Endogenous Retroviruses Derived from a Large White Pig and
RT Expression in Human Cells.";
RL J. Virol. 76:2714-2720(2002).
DR EMBL; AF435966; AAL87853.1; -; Genomic_DNA.
DR HSP; P03355; IMML.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00078; RVT; 1; 1.
DR PROSITE; PS50879; RNase H; 1.
SQ SEQUENCE 1139 AA; 127325 MW; 3B63C42301C04FAC CRC64;
Alignment Scores:
Pred. No.: 4.68e-309 Length: 1139
Score: 5383.50 Matches: 1054
Percent Similarity: 94.0% Conservative: 19
Best Local Similarity: 92.3% Mismatches: 52
Query Match: 36.8% Indels: 17
DB: 2 Gaps: 13
US-10-723-552-3 (1-8132) x Q8Q6U7_9GAMR (1-1139)
Qy 2347 GAAGAACAGTTGACTGGGAGTGGGACGGGTAAACCCACTCGTTCTGTCATACCTG--- 2403
Db 4 GluAspGlnLeuThrTrpGluLeuAspGlyGluProThrGlyPheTrpValIleLeuLys 23
Qy 2404 AGTCCCGACACCCCTCTTAGGTAGACATATTATGACCAAGATCGGAGCAAAATTTTC-- 2461
Db 24 Cys-ProSerThrLeuLeuArgValGlnThrTyraSpGlnAspGlyLysProLysPheLeu 43
Qy 2462 -TTTGAACAGGGAACCC---AGAAGTGTCTGCAATAACAAACCTATCATCTGTGTTGA 2517
Db 2462 -TTTGAACAGGGAACCC---AGAAGTGTCTGCAATAACAAACCTATCATCTGTGTTGA 2517
```

Db	43	upheGluProArgGluThrGlnLysCysLeu-GlnAsnThrAsnProIleThrValLeuT	63	Qy	3580	CTCTGGCCCTCCCTGACGTAAACCCCTTTACCCCTTATGTGGATGAGCGTAAAGGAG	3639
Qy	2518	CCCTCCCAATTAGATGACGAATATCGACTATATCTCTCCCTCTAGTAAAGCCTCATCAAAATA	2577	Db	419	laLeuAlaLeuProAspValLeuLysProPheThrLeuTyr-ValAspGluArgLysGlyV	439
Db	63	hrLeuGlnLeu---AspGluTyrArgLeuTyrSerProLeuValLysProAspGlnAsnI	82	Qy	3640	TAGCCCGGGAGAGTTTAAACCCAAACCCCTAGACCATGGAGAACCTGTCTCCCTACCTGT	3699
Qy	2578	TACAAATCTGTTGGACACAGTTTCCCAAGCCTGGCGAAGAACCGCAGGGATGGTTGG	2637	Db	439	alAlaArgGlyValLeuThrGlnIleLeuGlyProTyrArgArgProValAlaTyrLeuS	459
Db	82	leGlnPheTrpLeuGluGlnPheProLysAlaTrpAlaGluThrAlaGlyMetGlyLeuA	102	Qy	3700	CAAAAGAGCTCGATCTCTGTAGCAGTGTGGCCCATATGCTGAAGCTATCGCAGCTG	3759
Qy	2638	CAAAAGCAAGTTCCCCCAACAAGTATTCAACTGAAGCCAGTGCACACCCAGCTGTCAGTCA	2697	Db	459	erLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLysAlaIleAlaVal	479
Db	102	laLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaAlaProValSerVala	122	Qy	3760	TGGCCATPACTGGTCAAGGACGCTGCACAAATTTGACTTTGGGACAGATATTAACCTGTAATAG	3819
Qy	2698	GACAGTACCCCTTGAGTAAAGAAGCTCAAGAAGAAATTCGGCCGCATGTCCAAAGATTAA	2757	Db	479	alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleA	499
Db	122	rgGlnTyrLeuLeuSerLysGluAlaArgGluGlyIleGlyProHisValGlnArgLeuI	142	Qy	3820	CCCCCATGATGGGAGAAACATGTTCCGGCAGCCCCCAGACCCGATGATGATCAACACGCC	3879
Qy	2758	TCCAACAGGGCATCTAGTCTCTGTCCTCCATCTCCCTGGAATACTCCCTGCTACCGTTA	2817	Db	499	laProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrpMetThrAsnAlaA	519
Db	142	leGlnGlnGlyIleLeuValProValGlnSerProTyrAsnThrProLeuLeuProVala	162	Qy	3880	GCATGACCCACTATCAAGGCTGCTCTCACAGAGAGGGTCAGTTTCGCTCCACACGCCG	3939
Qy	2818	GAAGGCTGGGACTAATGACTATCGACAGTACAGGACTTCGAGAGAGTCAATAAACGGG	2877	Db	519	rgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPheAlaProProAlaA	539
Db	162	rgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgV	182	Qy	3940	CTCTCAACCTTGGCCACTCTTCTGCTGAAGAGACTGATGAACCACTGACTCATGATTGCC	3999
Qy	2878	TGCAGGATATACACCAACAGTCCCGAACCCCTTATAACCTCTGTGTCTCTCCACCCC	2937	Db	539	laLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProValThrHisAspCysH	559
Db	182	alGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuCysAlaLeuProProG	202	Qy	4000	ATCAACTATTGATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATA---CCGCTGA	4056
Qy	2938	AACGGAGCTGTATACAGTATTGACTTTAAGGATGCTTCTTCTGCTGAGATTACACC	2997	Db	559	is---LeuLeuIleGluGluThrGlyValArgLysAspLeuThrAspIleProProLeuT	578
Db	202	lnArgSerTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisP	222	Qy	4057	CTGAGAGAGTCAACCTGCTTCACTACACGGAAGCAGTATGTGGTGGAGGTAAAGGA	4116
Qy	2998	CCACTAGCAACCACTTTTGGCTTCGAAATGAGAGATCCAGGTACCGGAGAACCGGC	3057	Db	578	hrGlyLysMetLeuThrTrpPheThrAspGlySerSerTyrValValGluGlyLysSerM	598
Db	222	roThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyAlaGlyArgThrGlyG	242	Qy	4117	TGGCTGGGGCGGGGTGGTGGACGGGACCCGCGACGATCTGGCCGAGCAGCTGCGGAAG	4176
Qy	3058	AGCTCACCTGGACCGACTGCGCCCAAGGGTTTCAAGAACTCTCCGACCATCTTTGACGAAG	3117	Db	598	etAlaGlyProProValValThrGlyThrArgThrIleTrpAlaSerSerLeuProGluG	618
Db	242	lnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluA	262	Qy	4177	GAATTCAGACACAAAAGGCTGAGCTCATGCTCCCTCAGCGAAGCTTTCGGCTGCGCGAAG	4236
Qy	3118	CCCTACACAGAGCTGGCCACTTCAGGATCCAAACAACCTCAGGTGACCTCTCTCCAGT	3177	Db	618	lyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluG	638
Db	262	laLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnT	282	Qy	4237	GGAAATCCATAACATTTATACGACAGCAGGTATGCTTTGGACTGCACACCTACATG	4296
Qy	3178	ACGTGGATGACCTCTCTCGCGGAGCCACCAACAGGACTGCTTAGAAGGCACGAGG	3237	Db	638	lyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisG	658
Db	282	yrValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGly-----	299	Qy	4297	GGGCCATCTATAACAAAAGGGGTGCTTACCTCAGCAGGAGGGGAATAAAGAACAAAG	4356
Qy	3238	CACCTACTGCTGGAATTTGCTGACCTAGGCTACAGAGCTCTCTGCTAAGAAGGCCGAGATT	3297	Db	658	lyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysThrLysG	678
Db	300	--LeuLeuLeuGluLeuPheAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleC	319	Qy	4357	AGGAATTTCAAGCCTATTAGAACCGCTACATTTACAAAGAGCTAGCTATTATACACT	4416
Qy	3298	GCAGGAGAGGTAAACATCTTGGGGTACAGTTTG---CGGGACGGGACGGA---TGGC	3351	Db	678	luGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisC	698
Db	319	ysArgArgGluAlaThrGlnLeuGlyValGlnValCysGlyAlaGlyGlnSerAspTrpL	339	Qy	4417	GTCTCGACATCAGAAGCTTAAAGATCTCATATCCAGAGGAAACCCAGATCGCTGACCGGG	4476
Qy	3352	TGAGC---GAGGACCGAAGAAACTCTA---GTCCAGATACCGGCCCAACACACAGCA	3405	Db	698	ysProGlyHisGlnLysAlaLysAspProIleSerArgGlyAsnGlnMecAlaAspArgV	718
Db	339	euthrGlyLysAlaArgLysLysThrValGlnProLysIleGlyProProThrThrAlaL	359	Qy	4477	TTGCCAAGCAGGCGCCCGGGTGTAAACCTTCTGCTTATAATAGAAATGCCCAAGGCC	4536
Qy	3406	AACAA---ATCAGAGAGTGTTTTGGGGACA---GCTGGATTTTGAGACTGTGATCCCGG	3459	Db	718	alAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProMetIleGluThrProLysAlaP	738
Db	359	ysGlnValValArgGluPhePheGlyAlaGlnValGlyPheCysArgLeuTrpIleProG	379	Qy	4537	CAGAACCCAGACAGCATACACCTTAGAGACTGGCAAGAGATATAAAGATAGACCACT	4596
Qy	3460	GGTTTGGCACTTAGACGCCCACTCTACCGCTTAACCAAGAAAGAGGGAATTCCTCT	3519	Db	738	roGluProGlyArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLysIleAspGlnP	758
Db	379	lyPheAlaThrLeuAlaAProLeuTyrProLeuThrLysGluLysGlyGluPheSerT	399	Qy	4597	TCCTCAGACTCCGGAGGAGCCTGTATACCTCAGATGGGAGGAATCTCTCCCCACA	4656
Qy	3520	GGGCTCTGAGCAGCAGAGGCATTGATGCTATCAAAAGGCCCTCTCAGCGCACCTG	3579	Db	758	heSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIleLeuProHisL	778
Db	399	rpAlaLeuGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeuSerSerAlaProA	419				

```
QY 4657 AAGAAGGCTTAGAATAATGTCACACAGATACATCGTCAACCCAGCTAGGAACCTAAACACC 4716
Db |||
778 ySGluGlyLeuLysrValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisL 798
QY 4717 TGCAGCAGTTCGCTCAGAACATCCCTTATCATGTTCTGAGGCTTACCAGGAGTGCCTGACT 4776
Db |||
798 euGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGlyValAlaAspS 818
QY 4777 CGGTGGTCAACACATTGTGTGCTGCCAGCTGGTAAATGCTTAATCCCTCCAGAATGCCTC 4836
Db |||
818 erValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSerArgIleProp 838
QY 4837 CAGGGAAGAGACTAAGGGGAAGCCACCCAGGCGCTCACTGGGAAGTGGACTTCACCTGAGG 4896
Db |||
838 roGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluV 858
QY 4897 TAAAGCCCGCTAAATACGGAACAAATACCTATTGCTTTTGTAGACACCTTTTTCAGGAT 4956
Db |||
858 alLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyT 878
QY 4957 GGGTAGAGCTTATCCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAACCAATACTGG 5016
Db |||
878 rpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLysIleLeuG 898
QY 5017 AAGAAATTTTCCAGATTTCGSAATACCTAAGGTAATAGGCTCAGACAATGGTCCAGCTT 5076
Db |||
898 luGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsnGlyProAlap 918
QY 5077 TTGTTGCCCAGGTAAAGTCAGGAGCTGGCCAAAGATATTTGGGGATTGATTGGAACTGCATT 5136
Db |||
918 heValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisC 938
QY 5137 GTGCATACAGACCCCAAGCTCAGGACAGGTAGAGAGGATGAATAGAACCATTTAAAGAGA 5196
Db |||
938 ysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluT 958
QY 5197 CCCTTACTAAATTCACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCCTGCCCTTTG 5256
Db |||
958 hrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeuLeuProPheV 978
QY 5257 TGCTTTTGTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTTATGAATTACTCT 5316
Db |||
978 alLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuT 998
QY 5317 ACGGGGACCCCCCATTTGTAGAAATTCCTTCTGTACATAGTCTGACCTGCTGCTTTT 5376
Db |||
998 yrGlyGlyProProProLeuValGluIleAlaSerValHisSerAlaAspValLeuLeus 1018
QY 5377 CCCAGCCTTCTCTTAGGCTCAAGGCATTGAGTGGTGGAGACAACGAGCGTGGAGGC 5436
Db |||
1018 erGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysG 1038
QY 5437 AACTCCGGGAGCCTTACTCAGGAGGAGGAGACTTGCAGATCCCATCGTTTCCAAAGTGG 5496
Db |||
1038 lnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnValProHisArgPheGlnValG 1058
QY 5497 GAGATTCACTTACGTCAGCCACCCCTGAGGAAACCTCAGAGACTCGGTGAGAGGCC 5556
Db |||
1058 lyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyP 1078
QY 5557 CTTATCTCGTACTTTTGACACACCAACGGCTGTGAAGTCGAAGGAATCTCCACCTGGA 5616
Db |||
1078 roTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIleSerThrTrpI 1098
QY 5617 TCCATGCAATCCCAACCGCTTAAACCGCGCCACCTCCCGATTCCGGGGTGGAAAGCCGAAAGA 5676
Db |||
1098 leHisAlaSerHisValLysProAlaProProProAspSerGlyTrpLysAlaGluLysT 1118
QY 5677 CTGAAAATCCCTTAAGCTTCGCTCCATCGCGGTGCTTCTTACTCTGTCAATACCTCT 5736
Db |||
1118 hrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerValAsnAsnSerS 1138
QY 5737 CA 5738
```

Db |||
1138 er 1138

Search completed: February 14, 2006, 14:45:02
Job time : 2183.71 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 12:52:14 ; Search time 216.36 Seconds
(without alignments)
3302.856 Million cell updates/sec

Title: US-10-723-552-3
Perfect score: 14636
Sequence: 1 GCGTGTGTACGACTGTGGG.....CTGTTTGATCAAAAAAAA 8132

Scoring table:
BLOSOM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB_spool/US10723552/runat_14022006_125139_12833/app_query.fasta_1
-DB=A Genesep -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -ENDS=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02p
-USER=US10723552 @CGN 1 1 734 @runat_14022006_125139_12833 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Genesep 21:*
1: genesep1980s:*
2: genesep1990s:*
3: genesep2000s:*
4: genesep2001s:*
5: genesep2002s:*
6: genesep2003as:*
7: genesep2003bs:*
8: genesep2004s:*
9: genesep2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	6099.5	41.7	1194	AAW39272	AAW39272 Porcine r
2	6038	41.3	1145	AAW32097	AAW32097 Miniature
3	6038	41.3	1145	AAW73286	AAW73286 Retroviral
4	6038	41.3	1145	AAW73447	AAW73447 Swine ret
5	5338	36.5	1737	AAW10044	AAW10044 MMLV gag-
6	5237	35.8	1784	AAW94427	AAW94427 FeLV F6A
7	5229	35.7	1784	AAW05898	AAW05898 Gene prod
8	4889	33.4	1069	AAW73284	AAW73284 Defective
9	4332.5	29.6	1203	AAW81572	AAW81572 Mus dunni

10	3981	27.2	1196	2	AAW75189	Osteoindu
11	3980.5	27.2	1199	7	ADH76470	Murine le
12	3979.5	27.2	1197	4	AAW49468	Canine re
13	3969.5	27.1	1199	3	AAW12994	MLV rever
14	3953	27.0	1224	2	AAW17947	MOMLV pol
15	3850	26.2	744	8	ADS73450	Swine ret
16	3600	24.6	1079	2	AAW03163	MuLV reve
17	3534	24.1	666	8	ADS73449	Swine ret
18	3430	23.4	638	4	AAW32098	Miniature
19	3430	23.4	638	4	AAW70633	Porcine e
20	3430	23.4	638	4	AAW73287	Retroviral
21	3430	23.4	638	8	ADS73448	Swine ret
22	3425	23.4	638	4	AAW35114	PERV-C en
23	3412	23.3	638	9	ADY28028	Porcine e
24	3412	23.3	638	9	ADY28030	Porcine e
25	3412	23.3	638	9	ADY28036	Porcine e
26	3412	23.3	638	9	ADY28041	Porcine e
27	3412	23.3	638	9	ADY28038	Porcine e
28	3407	23.3	638	4	AAW70634	Porcine e
29	3318.5	22.7	678	9	ADY28032	Porcine e
30	3318.5	22.7	678	9	ADY28034	Porcine e
31	3214	22.0	1193	5	ABW79876	Spleen ne
32	3147.5	21.5	653	6	ABW82635	PERV enve
33	3135.5	21.4	653	6	ABW82636	PERV enve
34	2940.5	20.1	659	4	AAW35112	PERV-1-15
35	2879.5	19.7	661	9	ADY28039	Porcine e
36	2874.5	19.6	660	9	ADY28012	Porcine e
37	2871.5	19.6	660	2	ADY28004	Porcine e
38	2866.5	19.6	660	2	AAW85453	Pig endog
39	2866.5	19.6	660	4	AAW35113	PERV-A en
40	2866.5	19.6	660	4	AAW70632	Porcine e
41	2851	19.5	678	9	ADY28008	Porcine e
42	2849.5	19.5	660	9	ADY28014	Porcine e
43	2849.5	19.5	660	9	ADY28006	Porcine e
44	2835	19.4	678	9	ADY28010	Porcine e
45	2827	19.3	540	8	ADS73451	Swine ret

ALIGNMENTS

RESULT 1

AAW39272
ID AAW39272 standard; protein; 1194 AA.

XX
XX AAW39272;
XX AC
XX 27-AUG-2003 (revised)
DT 19-MAY-1998 (first entry)
XX
XX Porcine retrovirus POL protein.
XX
XX Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein; vaccine;
XX diagnosis; xenotransplantation; prophylactic; therapeutic.
XX Pig endogenous retrovirus.
XX
XX WO9740167-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-GB001087.
XX
XX 19-APR-1996; 96GB-00008164.
XX 10-FEB-1997; 97GB-00002668.
XX
XX (QONE-) Q-ONE BIOTECH LTD.
XX (IMUT-) IMUTRAN LTD.
XX Galbraith DN, Haworth C, Lees GM, Smith KT;
XX WPI; 1997-535851/49.
XX DR N-PSDB; AAW09700.
XX

PT Polynucleotide encoding porcine retrovirus expression product - useful to
 develop products for use in vaccines, diagnosis and xeno-transplantation.

PS Claim 6; Fig 3; 69pp; English.

XX This sequence represents the porcine retrovirus (PoRV) polymerase (POL)
 CC protein. This protein and other porcine retroviral proteins e.g. the
 CC virion core (GAG) and envelope (ENV) proteins can be used to develop
 CC viral vaccines, antisense nucleic acids, ribozymes and other antiviral
 CC agents. They can also be used in xeno-transplantation technology and as
 CC diagnostic tools. (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 1194 AA;

Alignment Scores:

Pred. No.: 0 Length: 1194
 Score: 6099.50 Matches: 1154
 Percent Similarity: 98.5% Conservative: 19
 Best Local Similarity: 96.9% Mismatches: 17
 Query Match: 41.7% Indels: 1
 DB: 2 Gaps: 1

US-10-723-552-3 (1-8132) x AAW39272 (1-1194)

QY	2160	GGCAGACGGGTTTCGACACCCCTCCCGAGCCCGAGGTAACCTTTGAAGGTGGAGGGCAA	2219
DB	1	GIYARGGlySerAspProLeuProGluProArgValThrLeuLysValGluGlyGln	20
QY	2220	CCAGTTGAGTTCCTGGTTGATACCGGAGCGAAACATTCAGTCTACTACAGCCATTAGGA	2279
DB	21	ProValGluPheLeuValAspThrGlyAlaGluHisSerValLeuLeuGlnProLeuGly	40
QY	2280	AACTAAAGATAAAAATCTCGGTGATGGGTGCCACAGGCAACACAGATTCATGG	2339
DB	41	LysLeuLysGluLysSerTrpValMetGlyAlaThrGlyGlnArgGlnTrpProTrp	60
QY	2340	ACTACCGAAGACAGTTGAGTGGAGTGGGACGGTACCCACTCGTTTCGTGTCATA	2399
DB	61	ThrThrArgargThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValIle	80
QY	2400	CCTGAGTGCACACCCCTCTTAGGTAGACACTTATTGACCAAGATGGGAGCACAAAT	2459
DB	81	ProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle	100
QY	2460	TCCTTTGAACAAGGAAAACAGAGCTGTCGAAATAACAACTCATCTACTGTGTGACC	2519
DB	101	SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleThrValLeuThr	120
QY	2520	CTCCAATTAGATCAGCAATATCGACTATCTCTCCCTAGTAAAGCTGTATCAAAATATA	2579
DB	121	LeuGlnLeuAspAspGluTrpArgLeuTrpSerProGlnValLysProAspGlnAspIle	140
QY	2580	CAATTCTGGTTGGAACAGTTTCCCAAGCCTGGGCGAGAAACCGCAGGATGGGTTTGGCA	2639
DB	141	GlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuAla	160
QY	2640	AAGCAGTTCCCCCACAAGTTATTCAACTGAAGCGCAGTCCACACAGTGTCACTCAGA	2699
DB	161	LysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValArg	180
QY	2700	CAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCCGATGTCCAAAGATTAATC	2759
DB	181	GlnTrpProLeuSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle	200
QY	2760	CAACAGGCACTCCTAGTTCCTGTCCTCAATCTCCCTGGAAATACTCCCTGTCTACCGGTTAGA	2819
DB	201	GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValArg	220
QY	2820	AAGCTGGGACTAATGACTATCGACGATCAGAGCTTGAGAGAGGTCAATAACGGGTG	2879
DB	221	LysProGlyThrAsnAspTrpArgProValGlnAspLeuArgGluValAsnLysArgVal	240
QY	2880	CAGGATATACACCAACAGTCCCGAACCCCTTATAACCTCTGTGTCTCTCCACCCCAA	2939

DB	241	GlnAspIleHisProThrValProAsnProTrpAsnLeuLeuSerAlaLeuProProGlu	260
QY	2940	CGGAGCTGTATACAGTATTGGACTTAAAGGATGCCTTCTTCCTGCTGAGATTACACCCC	2999
DB	261	ArgAsnTrpTrpThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisPro	280
QY	3000	ACTAGCAACCACTTTTTCCTTCGAATGGAGATCCAGGTACGGGAAGAACCGGCGCAG	3059
DB	281	ThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln	300
QY	3060	CTCACCTGACCCGACTGCCCAAGGGTTCAAGAACTCCCGACCATCTTTTGACGAAGCC	3119
DB	301	LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluAla	320
QY	3120	CTACACAGAGACTGGCCAACTTCAGGATCCAAACACCTCAGTGACCTCTCTCCAGTAC	3179
DB	321	LeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnTrp	340
QY	3180	GTGGATGACTCTCTCTGCGGGGAGCCACCAACAGAGACTGTCTAGAAGGCACGAAGGCA	3239
DB	341	ValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysAla	360
QY	3240	CTACTCTCGAATTTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCACGATTTC	3299
DB	361	LeuLeuLeuGluLeuSerAspLeuGlyTrpArgAlaSerAlaLysLysAlaGlnIleCys	380
QY	3300	AGAGAGAGGTAACTACTTGGGGTACAGTTTCGCGGAGCGGCGAGCGATGGCTGAGGAG	3359
DB	381	ArgArgGluValThrTrpLeuGlyTrpSerLeuArgGlyGlyGlnArgTrpLeuThrGlu	400
QY	3360	GCACGGAAGAAACTGTAGTCCAGATACCGGCCCAACACACACCAAAACAAATGAGAGAG	3419
DB	401	AlaArgLysLysThrValGlnIleProAlaProThrThrAlaLysGlnValArgGlu	420
QY	3420	TTTTTGGGACAGCTGGATTTTCGAGACTGTGGATCCCGGGTTTCGCGACCTTAGCAGCC	3479
DB	421	PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAla	440
QY	3480	CACTCTACCCGTAACCAAGAAAGGGGAATTCCTCTGGGCTCTCTGAGCACCAAGAG	3539
DB	441	ProLeuTrpProLeuThrLysGluLysGlyPheSerTrpAlaProGlnHisGlnLys	460
QY	3540	GCATTTGATGCTATCAAAAGGCCCTGTGAGCGACCTGTCTGCGCCCTCCCTGACGTA	3599
DB	461	AlaPheAspAlaLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspVal	480
QY	3600	ACTAAACCCCTTTTACCTTTTATGTGGATGAGCGTAAAGGGAGTAGCCCGGGAGTTTAAAC	3659
DB	481	ThrLysProPheThrLeuTrpValAspGluArgLysGlyValAlaArgGlyValLeuThr	500
QY	3660	CAACCCCTAGGACCATGGAGAGACCTGTGCGCTACTCTGCAAGAAGCTCGATCTCTGTA	3719
DB	501	GlnThrLeuGlyProTrpArgProValAlaTrpLeuSerLysLysLeuAspProVal	520
QY	3720	GCAGTGGTTGGCCCATATGCTGAGGCTATCGCAGCTGTGGCCATCTCTGCAAGGAC	3779
DB	521	AlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaLeuValLysAsp	540
QY	3780	GCTGCAAAATTTGACTTTGGACAGATAAATCTGTATAGCCCCCCCCCATTCGATTGGAGAAC	3839
DB	541	AlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluAsn	560
QY	3840	ATCGTTTCGGCAGCCCCCAGACCGATGGATGACCAACCGCGGATGACCCACTATCAAGC	3899
DB	561	IleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTrpGlnSer	580
QY	3900	CTGCTTCTCACAGAGAGGTTCAGTTCGCTCCACACGCGCTCTCAACCTCGCACTCTT	3959
DB	581	LeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu	600
QY	3960	CTGCTTGAAGAGACTGATGAACCAAGTCACTGATTGCGCATCAACTATTGATTGAGGAG	4019

Db 601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuLeuGluGlu 620
QY 4020 ACTGGGTCGCAAGACCTTACAGACATACCCTGACTGAGAGAGTGCTAACTGGTTC 4079
Db 621 ThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpPhe 640
QY 4080 ACTGACGGAGCAGCTATGTGGTGGAGGTAAAGAGATGGCTGGCGCGGTGGTGGAC 4139
Db 641 ThrAspGlySerTyrValValGluGlyLysArgMetAlaGlyAlaValValAsp 660
QY 4140 GGCACCCGACACGATCTGGCGCACGACCTGCCGGAAGGAACCTTACGACACAAAAGCTGAG 4199
Db 661 GlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaGlu 680
QY 4200 CTCATGGCCCTCAGCAAGCTTTGCGGCTGGCGAAGGAAATCCATAAACATTTATAGC 4259
Db 681 LeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlySerIleAsnIleTyrThr 700
QY 4260 GACAGCAGGTATGCCCTTTGGGACTGCACAGTCATACATGGGGCCATCTATAACAAAGGGGG 4319
Db 701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgGly 720
QY 4320 TTGCTTACCTCAGCAGGGGGAATAAAGAACAAAGAGGMAATCTTAGCCCTATTAGAA 4379
Db 721 LeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuGlu 740
QY 4380 GCGCTACATTTACCAAAAGGCTAGCTATTATACACTGCTCGACATCAGAAAGCTAAA 4439
Db 741 AlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaLys 760
QY 4440 GATCTCATATCCAGAGAAACAGATGGCTGACCGGGTTGCCAAGCAGCAGCCAGGT 4499
Db 761 AspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaGlnAla 780
QY 4500 GTTAACCTTCTGCTATATAAGAAATGCCAAGCCCCAGAACCCAGACGACAGTACACC 4559
Db 781 ValAsnLeuLeuProIleLeuGluThrProLysAlaProGluProArgGlnTyrThr 800
QY 4560 CTAGAGACTGGCAAGAGATAAAAAGATAGACCACTTCTCAGACTCCGGAAGGGACC 4619
Db 801 LeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyThr 820
QY 4620 TGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAAAGAGGGTTAGAAATGTCCAA 4679
Db 821 CysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGluTyrValGln 840
QY 4680 CAGATACATCGTCTAACCCACTAGCACTAAACACCTGCAGCAGTGTGTCAGACATCC 4739
Db 841 GlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrSer 860
QY 4740 CCTTATCATGTTCTGAGGCTACAGGAGTGGCTGACTCGGTGCTCAACATTTGTGCCC 4799
Db 861 ProTyrHisValLeuArgLeuProGlyValAlaAspSerValValLysHisCysValPro 880
QY 4800 TGCCAGCTGGTAAATGCTTAATCCTTCCAGAATGCTTCCAGGAAGAGACTAAGGGGAAGC 4859
Db 881 CysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
QY 4860 CACCAGGCGCTCAGTGGGAAGTGGACTTACCTAGGTAAACCGCGCTAAATACCGAAAC 4919
Db 901 HisProGlyAlaHisIleTrpGluValAspPheThrGluValLysProAlaLysTyrGlyAsn 920
QY 4920 AAATACCTATTGCTTTTCTAGACACCTTTTCCAGGATGGGTAGAGCTTATCTCTACTAAG 4979
Db 921 LysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyrProThrLys 940
QY 4980 AAAGAGACTTCAACCGTGGTGGCTAAAAAATACTCGAAGAAATTTTCCAGATTTGGA 5039
Db 941 LysGluThrSerThrValValAlaLysLysIleLeuGluGluIlePheProArgPheGly 960
QY 5040 ATACCTAGGTAAATAGGTCAGACAAATGTCAGCTTTTGTGCCCAGGTAAAGTCAGGA 5099
Db 961 IleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnGly 980

QY 5100 CTGGCCAGATATTGGGGATTGATTGGAAACTGCATTGTGCATACAGACCCCAAGCTCA 5159
Db 981 LeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1000
QY 5160 GGACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCCCTTAATAATGACCGGGAG 5219
Db 1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrGlu 1020
QY 5220 ACTGGGTTAATGATTGGATAGCTCTCCCTGCCCTTTGTGCTTTTATAGGTTAGGACACC 5279
Db 1021 ThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnThr 1040
QY 5280 CTGGACACTTTGGGCTGACCCCTTATGAATTACTCTAGGGGGAGACCCCTCATTTGGTA 5339
Db 1041 ProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrGlyProProLeuAla 1060
QY 5340 GAAATTTGCTTGTATCATAGTGTGCTGCTGCTTTCCAGCCTTTTCTCTAGGCTC 5399
Db 1061 GluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeu 1080
QY 5400 RAGGCACCTTACGTGGGTGAGACAACGAGCGTGGAGCACTCCGGGAGGCTACTCAGGA 5459
Db 1081 LysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSer--- 1099
QY 5460 GGAGGAGACTTCACATCCACATCGTTTCCAAGTGGGAGATTTCAGTCTACGTTAGACGC 5519
Db 1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrValArgArg 1119
QY 5520 CACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGGCCCTTATCTCGTACTTTTGACCACA 5579
Db 1120 HisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeuValLeuLeuThrThr 1139
QY 5580 CCAACGGCTGCAAACTCGAAGGAATCTCCACCTGGATCCATGCTCCACCGTAAACCG 5639
Db 1140 ProThrAlaValLysValGluGlyIleProThrTrpIleHisAlaSerHisValLysPro 1159
QY 5640 GCGCACCTCCCGATTTCGGGGTGGAAAGCCGAAAGACTGAAAAATCCCTTAAGCTTCGC 5699
Db 1160 AlaProProProAspSerGlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArg 1179
QY 5700 CTCATCGGTGGTTCCTTACTCTCTCAATAAC 5732
Db 1180 LeuHisArgLeuValProTyrSerAsnAsnAsn 1190
RESULT 2
AAW32097
ID AAW32097 standard; protein; 1145 AA.
XX AAW32097;
XX 27-AUG-2003 (revised)
DT 09-FEB-1998 (first entry)
XX Miniature swine retrovirus POL protein.
XX Retrovirus; porcine; POL protein; xenotransplantation; infectious;
XX provirus; organ transplant; donor; activated virus; PCR.
XX Pig endogenous retrovirus.
FH Key Location/Qualifiers
FT Protein 1..1145
FT /label= POL_protein
XX WO9721836-A1.
XX 19-JUN-1997.
XX 13-DEC-1996; 96WO-US019680.
XX 14-DEC-1995; 95US-00572645.
XX

PA	(GEO) GEN HOSPITAL CORP.	PA
XX	Fishman JA;	XX
XX	WPI; 1997-332804/30.	XX
DR	N-PSDB; AAN74884.	DR
XX	New nucleic acid from porcine retroviruses - used for detecting viruses	XX
PT	in transplant or other tissue and for assessing risk of transmitting	PT
PT	infection to graft recipient.	PT
XX	Claim 22; Fig 3; 128pp; English.	XX
PS	This is a porcine retrovirus from miniature swine containing the coding	PS
XX	region for a putative viral POL protein. This sequence and PCR fragments	XX
CC	generated from the sequence (see A174812-174882) could be used to screen	CC
CC	organs for porcine retroviruses prior to xenotransplantation.	CC
CC	Transplantation can increase the likelihood of retroviral activation if	CC
CC	intact and infectious proviruses are present. The porcine retroviral	CC
CC	sequence can be used to generate probes to determine the level (e.g. copy	CC
CC	number) of intact (i.e. potentially replicating) porcine provirus	CC
CC	sequences in a strain of xenograft transplantation donors. It can be used	CC
CC	to detect mutations, genetic lesions or viral recombinants and also to	CC
CC	determine the histological localisation of activated retrovirus. Using	CC
CC	Polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear	CC
CC	cells, infectivity titration and susceptibility testing can be performed.	CC
CC	Ultimately animal donors without intact porcine retroviral sequences or a	CC
CC	lower copy number of viral elements could be selected. (Updated on 27-AUG	CC
CC	-2003 to correct OS field.)	CC
XX	Sequence 1145 AA;	XX
SQ		SQ
Alignment Scores:		
Pred. No.:	0	Length: 1145
Score:	6038.00	Matches: 1145
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0
Query Match:	41.3%	Indels: 0
DB:	2	Gaps: 0
US-10-723-552-3 (1-8132) x AAW32097 (1-1145)		
QY	2307 ATGGGTGCCACAGGCAACACAGTATCCATGGACTACCGAAGAACAGTTGACTTGGGA	2366
DB	1 MetGlyAlaThrGlyGlnGlnGlnTyrProTyrThrThrArgArgThrValAspLeuGly	20
QY	2367 GTGGGACGGGTAAACCCACTCGTTTCCTGTCATACCTGAGTGCACGACCCCTCTTAGGT	2426
DB	21 ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly	40
QY	2427 AGAGACTTATGGACCAAGATGGGAGCACAAATTTCTTTTCACAGGGAACCCAGAGTG	2486
DB	41 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGlnGlnGlyLysProGluVal	60
QY	2487 TCTGCAATTAACAAACCTATCACTGTGTTCAGCCCTCCAATTAGATGACCAATATCGACTA	2546
DB	61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu	80
QY	2547 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATTCGTGTTGAAACAGTTTCCCCAA	2606
DB	81 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln	100
QY	2607 GCCTGGGCAGAAACCCGAGGATGGTTGGCAAGCAAGTTCCTCCCAAGTATTTCAA	2666
DB	101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln	120
QY	2667 CTGAAGCCAGTGCCACACAGGTGTCAGTCAGACAGTACCCTTGTAGTAAAGAGCTCAA	2726
DB	121 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln	140
QY	2727 GAAGGAATTCGGCCGATGTCCTCAAGATTAATCAACAGGGCATCTAGTCTCTGTCCAA	2786
DB	141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln	160

QY	2787 TCTCCCTCGAATACTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCA	2846
DB	161 SerProTyrPAsnThrProLeuLeuProValAlaGlyLysProGlyThrAsnAspTyrArgPro	180
QY	2847 GTACAGGACTTGAGAGAGGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAAC	2906
DB	181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn	200
QY	2907 CTTATAACCTCTTGTGTGCTCTCCACCCCAACCGAGCTGGTATACAGTATTGGACTTA	2966
DB	201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTyrThrValLeuAspLeu	220
QY	2967 AAGGATGCCTTCTTCTGCTGAGATTACACCCCACTAGCAACCACTTTTGGCTTCGAA	3026
DB	221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu	240
QY	3027 TGGAGAGATCCAGGTACCGGGAAGAACCGGGCAGCTCACCTGGACCCGACTGCCCCAAAGG	3086
DB	241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTyrThrArgLeuProGlnGly	260
QY	3087 TTCAGAAGCTCCCGGACCATCTTTGACGAAGCCCTACACAGAGACTGGCCAACTTCAGG	3146
DB	261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg	280
QY	3147 ATCCAAACCCCTCAGGTGACCCCTCCTCCAGTAGCTGATGACCTGCTTCTGGCGGGAGCC	3206
DB	281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla	300
QY	3207 ACCAAACAGGACTGCTTAGAAGGCACGAAGGCACACTACTGCTGAATGTCTGACCTAGGC	3266
DB	301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly	320
QY	3267 TACAGAGCTCTGCTTAAGAAGCCCAAGATTGCGAGAGAGAGTAAACATATTGGGGTAC	3326
DB	321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgGluValThrTyrLeuGlyTyr	340
QY	3327 AGTTTTCGGGACGGCGAGCGATGGCTGACGGAGGCGACGGAAGAAACTGTAGTCCAGATA	3386
DB	341 SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle	360
QY	3387 CCGGCCCCAACACACAGCCAAACAAATGAGAGAGTTTTTGGGGACACCTGGATTTTCAGA	3446
DB	361 ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg	380
QY	3447 CTGTGATCCCGGGTTTCGACCTTAGCAGCCCGACTACCCGCTACCAACCAAGAAAA	3506
DB	381 LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLys	400
QY	3507 GGGGAATTCCTCTGGGCTCTGAGCACCAAGAGGCAATTTGATGCTATCAAAAAGGCCCTG	3566
DB	401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu	420
QY	3567 CTGAGCGCACCTGCTCTGGCCCTCCTGACGTAACTAAACCCCTTTACCTTTATGTGGAT	3626
DB	421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp	440
QY	3627 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTAGGACCATGGAGAACCT	3686
DB	441 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro	460
QY	3687 GTCCCTTACTCTCAAGAGACTCCGATCCTGTAGCCAGTGGTTGGCCCATATATCCCTGAAG	3746
DB	461 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLys	480
QY	3747 GCTATCGCAGCTGTGGCCATCTGCTCAAGGAGCGCTGACAAATTTGACTTTGGACAGAAT	3806
DB	481 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn	500
QY	3807 ATAACGTGAATAGCCCCCATGATTGGAGAACATCGTTGGCAGGCCCCCGACCGCATGG	3866
DB	501 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp	520

QY 3867 ATGACCAAGCCCGCATGACCCACTATCAAGCGCTGCTTCTCAGAGAGGGTCAAGCTTC 3926
DB 521 MetThrAsnAlaargMetThrHisTyrglnSerLeuLeuLeuThrGluArgValThrPhe 540
QY 3927 GCTCCACAGCGCTCTCAACCCCTGCCACTCTTCTGCTCAAGAGACTGATGAACAGTG 3986
DB 541 AlaProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 3987 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 4046
DB 561 ThrHisAspCysHisGlnLeuLeuLeuGluThrGlyValargLysAspLeuThrAsp 580
QY 4047 ATACCCCTGACTGGAGAGTCTAACTGTTTCACTGACGGAAGCAGCTATGTGTGAA 4106
DB 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerTyrrValValGlu 600
QY 4107 GGTAAAGAGGATGCTGGGGCGGCTGTGCGAGCGGACCGCACGATCTGGGCCACGAC 4166
DB 601 GlyLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
QY 4167 CTCCCGAAGGAATCTCAGCACAAAAGGCTGAGCTCATGCGCCCTCAGCAAGCTTTGCGG 4226
DB 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 4227 CTGGCCGAAGGGAATCCATAACATTTATACGACAGCAGGTATGCTTTGCGACTGCA 4286
DB 641 LeuAlaGluGlyLysSerIleAsnIleTyrrThrAspSerArgTyrrAlaPheAlaThrAla 660
QY 4287 CAGTCATGCGGCCCATCTATAACAAAGGGGTGCTTACCTCAGCAGGGGGGAATA 4346
DB 661 HisValHisGlyAlaIleTyrrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 4347 AAGAACAAAGAGGAAATCTTAAGCCTATTAGAAGCGCTACATTTACCAAAAAGGCTAGCT 4406
DB 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
QY 4407 ATTATACATGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACCAAGTG 4466
DB 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
QY 4467 GCTGACGGGTGCCAAGCAGCGAGCCAGGGTGTAACTTCTGCTATATAGAAATG 4526
DB 721 AlaAspArgValAlaLysGlnAlaGlnGlyValAsnLeuLeuProIleIleGluMet 740
QY 4527 CCCAAAGCCCGCAAGCCAGACAGTACACCTAGAAGACTGGCAAGAGATAAAAAAG 4586
DB 741 ProLysAlaProGluProArgArgGlnTyrrThrLeuGluAspTrpGlnGluIleLysLys 760
QY 4587 ATAGACCACTTCTTGAGACTCCGGAAGGACTGCTATACCTCAGATGGGAAGAAATC 4646
DB 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrrThrSerAspGlyLysGluIle 780
QY 4647 CTGCCCCACAAAGAGGGTTAGATATGTCACAGATACATCGTCTAACCCACCTAGGA 4706
DB 781 LeuProHisLysGluGlyLeuGluTyrrValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 4707 ACTAAACACTCGCAGCAGTTGTCAGAACATCCCTTATCATGTCTTGAGGCTACCAGGA 4766
DB 801 ThrLysHisLeuGlnGlnLeuValargThrSerProTyrrHisValLeuArgLeuProGly 820
QY 4767 GTGGCTGACTCGGTGTCAAAATGTTGTGCGCTGCCAGCTGTTTAATGCTAATCTTCC 4826
DB 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY 4827 AGAATCCCTCCAGGGAAGACTAAGGGAAGCCACCGGCGCTCACTGGGAAGTGGAC 4886
DB 841 ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
QY 4887 TTCACTGAGGTAAAGCCGCTAAATAGCGAAACAAATACCTATTGTTTTCTGACACACC 4946
DB 861 PheThrGluValLysProAlaLysTyrrGlyAsnLysTyrrLeuLeuValPheValAspThr 880
QY 4947 TTTTCAGGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGTGGCTAAA 5006

DB 881 PheSerGlyTrpValGluAlaTyrrProThrLysLysGluThrSerThrValValAlaLys 900
QY 5007 AAAATACTGGAGAAATTTTCCAAAGATTGGTAATACCTAAGCTAATAGGGTCAGACAAT 5066
DB 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
QY 5067 GGTCCAGCTTTTGTCTCCCAAGTAAGTCAGGGACTGCCCCAAGATATTGGGGATTGATTGG 5126
DB 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY 5127 AAATCTGATTGTGCATACAGACCCCAAGCTCAGGACAGCTAGAGAGATGAATAGAAC 5186
DB 941 LysLeuHisCysAlaTyrrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY 5187 ATTAAAGAGACCTTACTAAATTGACCGGGAGACTGGCGTTAATGATTGGTAGCTCTC 5246
DB 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
QY 5247 CTGCCCTTTGTCTTTTAGGGTTAGGAACACCCCTGGACAGCTTGGGCTGACCCCTAT 5306
DB 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrr 1000
QY 5307 GAATTAATCTACGGGGGACCCCCCATTTGGTAGAAATTCCTTCTACATAGTGTGAC 5366
DB 1001 GluLeuLeuTyrrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
QY 5367 GTGCTGCTTCCAGCGCTTGTCTCTAGCTCAGGCACTTGGTGGTGAGACACGA 5426
DB 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY 5427 GCGTGGAGCAACTCCGGGAGCGCTACTCAGAGGAGGAGACTTGCAGATCCACATCGT 5486
DB 1041 AlaTrpArgGlnLeuArgGluAlaTyrrSerGlyGlyAspLeuGlnIleProHisArg 1060
QY 5487 TTCCAAGTGGAGATTCACTTACGTTAGACCCCGCTGCAGGAAACCTCCGAGACTCGG 5546
DB 1061 PheGlnValGlyAspSerValTyrrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
QY 5547 TCGAAGGCGCTTATCTCGTACTTTTGACCACACCGCTGTGAAGTCGAGGAATC 5606
DB 1081 TrpLysGlyProTyrrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
QY 5607 TCCACTGGATCCATCCCATCCAGCTTAAACCGGCGCACCTCCCGATTCCGGGTGAAA 5666
DB 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1120
QY 5667 GCCGAAAAGACTGAAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTC 5726
DB 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrrSerVal 1140
QY 5727 ANTAACCTCTCAGAC 5741
DB 1141 AsnAsnLeuSerAsp 1145
RESULT 3
ID AAB73286
XX AAB73286 standard; protein; 1145 AA.
AC AAB73286;
XX
DT 23-MAY-2001 (first entry)
XX
DE Retroviral protein #2 found in miniature swine.
XX Retrovirus; graft transplantation; xenotransplantation; miniature swine.
XX Unidentified.
XX
PN US6190861-B1.
XX
PD 20-FEB-2001.
XX

PF	13-DEC-1996;	96US-00766528.	
XX			
PR	14-DEC-1995;	95US-00572645.	
XX			
PA	(GEOH) GEN HOSPITAL CORP.		
XX			
PI	Fishman JA;		
XX			
DR	WPI; 2001-256211/26.		
DR	N-PSDB; AAF77727.		
XX			
PT	Assessing risk of endogenous retroviruses in clinical practice and in		
PT	xenotransplantation, comprises using probe sequences derived from swine		
PT	or miniature swine retroviral genome.		
XX			
PS	Disclosure; Fig 3; 127pp; English.		
XX			
CC	The present invention relates to a method for screening a cell or tissue		
CC	for the presence or expression of a retrovirus (RV), comprising		
CC	contacting a target nucleic acid from the cell or tissue with a second		
CC	nucleic acid from the present invention (e.g. AAF77727 or a fragment		
CC	thereof). The method is useful for RV detection and to assess graft		
CC	transplantation risk. Screening of animals allows the elimination of		
CC	donors with active replication of known viruses. Inactive proviruses can		
CC	be detected and inactivated, allowing identification and elimination of		
CC	potential human pathogens derived from swine in a manner not possible in		
CC	the outbred human organ donor population and is important to the		
CC	development of human xenotransplantation		
XX			
SQ	Sequence 1145 AA;		
Alignment Scores:			
Pred. No.:	0	Length:	1145
Score:	6038.00	Matches:	1145
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	41.3%	Indels:	0
DB:	4	Gaps:	0
US-10-723-552-3 (1-8132) x AAB73286 (1-1145)			
QY	2307	ATGGGTGCCACAGCGCAACACAGTATCCATGGACTACCCGAAAGACAGTGTGACTTGGGA	2366
DB	1	MetGlyAlaThrGlyGlnGlnTyrProTyrThrArgArgThrValAspLeuGly	20
QY	2367	GTGGGACGGTAAACCACTCGTTCTGGTCATACCTGAGTGCCACGCCCTCTTAGGT	2426
DB	21	ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly	40
QY	2427	AGAGACTTATTGACCAAGATGGGAGCACAAATTTCTTTGAACAAGSGRAACCAAGATG	2486
DB	41	ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal	60
QY	2487	TCTGCAATAACAAACCTATCACTGTGTGACCTCCAAATTAGATGACCAATATCGACTA	2546
DB	61	SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu	80
QY	2547	TACTCTCCCTAGTAAAGCTGATCAAAATATACAAATTCGTGTTGGAAACAGTTTCCCAA	2606
DB	81	TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln	100
QY	2607	GCCTGGGCAGAAACCCAGGGATGGTTGGCAAGCAAGTTCCTCCCAAGATTATTCAA	2666
DB	101	AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProGlnValIleGln	120
QY	2667	CTGAAGCCAGTCCCAACACAGTGTGAGTCAGACAGTACCCCTTGAGTAAAGAGCTCAA	2726
DB	121	LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln	140
QY	2727	GAAGGAATTCCGGCGGATGCCAAAGATTAAATCCAAACAGGGCATCCTAGTCTCTGCCAA	2786
DB	141	GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln	160

QY	2787	TCTCCCTGGAATACCTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCA	2846
DB	161	SerProTyrAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro	180
QY	2847	GTACAGGACTTGAGAGAGGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAAC	2906
DB	181	ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn	200
QY	2907	CCTTATAACCTCTGTGTGCTCTCCACACCCCAACGGAGCTGGTATACAGTATTGGACTTA	2966
DB	201	ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu	220
QY	2967	AAGGATGCTCTTCTTCCTGCTGAGATTACACCCCACTAGCAACCACTTTTTCCTTCGAA	3026
DB	221	LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu	240
QY	3027	TGAGAGATCCAGGTACGGGAAGAACCCGGGAGCTCACTGGACCCGACTGCCCCCAAGGG	3086
DB	241	TyrArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly	260
QY	3087	TTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTCAGG	3146
DB	261	PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg	280
QY	3147	ATCCAAACCCCTCAGGTGACCCCTCCTCCAGTACCTGGATGACCTGCTTCTGGCGGAGCC	3206
DB	281	IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla	300
QY	3207	ACCAACACGAGCTGTTCAGAGGCACGAAGGCACACTACTCTGGAATTGCTGACCTAGGC	3266
DB	301	ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAspLeuGly	320
QY	3267	TACAGAGCTCTGTAAAGAGCCCAAGATTTCCAGAGAGAGTACATCTTGGGGTAC	3326
DB	321	TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr	340
QY	3327	AGTTTCGGGACGGGACGGATGGCTGACGGAGGCACGGAAGAAAACTGTAGTCCAGATA	3386
DB	341	SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysThrValValGlnIle	360
QY	3387	CGGGCCCCAACCAAGCCAAATGAGAGAGTTTTTGGGACAGCTGGATTTTCGAGA	3446
DB	361	ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg	380
QY	3447	CTGTGATCCCGGGTTTCGGACCTTAGCAGCCCACTTACCCGCTAACCAAGAAAAA	3506
DB	381	LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLys	400
QY	3507	GGGGAATTCCTCTGGGCTCTGAGCACCAAGAGCAATTTGATGCTATCAAAAAGGCCCTG	3566
DB	401	GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu	420
QY	3567	CTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTTACCTTTATGTGGAT	3626
DB	421	LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp	440
QY	3627	GAGCGTAAAGGAGTAGCCCGGGAGTTTAAACCAACCTTAGGACCATGGAGAAGACCT	3686
DB	441	GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro	460
QY	3687	GTGCGCTACTGTCAAGAAAGCTCGATCCTGTAGCCAGTGGTTGGGCCCATATGCTCGAAG	3746
DB	461	ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLys	480
QY	3747	GCTATCCGAGCTGTGGCCATCTGCTCAAGGACGCTGACAAATTGACTTTTGGACAGAAT	3806
DB	481	AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn	500
QY	3807	ATTAACGTGAATAGCCCCCATGCAATGGAGACATCGTTCCGACGCCCCAGACCCGATGG	3866
DB	501	IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp	520
QY	3867	ATGACCAACGCCCGCATGATGACCCCACTATCAAGCCTCTCTCTCACAGAGAGGTTCACGTT	3926

Db 521 MetThrAsnAlaArgMetThrHisIleTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540
QY 3927 GCTCCACCGCCCTCTCAACCCCTGCGCAGCTCTTCTGCTCAAGAGAGCTGATGAACAGTG 3986
Db 541 AlaProProAlaAlaLeuLeuProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 3987 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 4046
Db 561 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580
QY 4047 ATACCCCTGACTCGAGAGTGCTAACTGGTTCACTGACGGAAGAGCTATGTGTGGAA 4106
Db 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
QY 4107 GGTAAAGAGATGCTCGGGCGCGGTGGTGAGCGGACCGCACGATCTGGGCCACGAC 4166
Db 601 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
QY 4167 CTGCCCGAAGGAACTTCAGCACAAAAGGCTGAGCTCATGGCCCTCACGCAAGCTTTGCGG 4226
Db 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 4227 CTGGCCGAAGGGAATCCATAACATTATACGGAAGAGAGTATGCTTTGCGACTGCA 4286
Db 641 LeuAlaGluGlyThrSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
QY 4287 CAGGTACATGGGCGCCATCTATAACAAAGGGGTGCTTACCTCAGCAGCGGGGGAATA 4346
Db 661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 4347 AAGAAACAAAGAGAAATTTCTAAGCCATTAGAGCCGTATACATTTACCAAAAAGGCTAGCT 4406
Db 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
QY 4407 ATTATACATGCTCTGGACATCGAAAGCTAAAGATCTCATATTCGAGAGAAACGAGTG 4466
Db 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
QY 4467 GCTGACCGGGTTCGCAAGCAGGCGCCAGGGTGTAACTTCTGCTTATATAGAAATG 4526
Db 721 AlaAspArgValAlaLysGlnAlaIleGlnGlyValAsnLeuLeuProIleIleGluMet 740
QY 4527 CCCAAAGCCCGAAGCCAGACGACAGTACACCCCTAGAGACTGGCAAGAGATAAAAAAG 4586
Db 741 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
QY 4587 ATAGACCAAGTCTCTGAGACTCCGGAAGGAGCTGCTATACCTCAGATGGGAAGGAATC 4646
Db 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
QY 4647 CTGCCCCCAAGAGAGGGTTAGAAATATGTCACAGATACATCGTCAACCCACCTAGGA 4706
Db 781 LeuProHisLysGlyGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 4707 ACTAAACACCTCGCAGCAGTTGCTCAGAACATCCCTTATCATGTTCTGAGGCTACCAGA 4766
Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
QY 4767 GTGGCTGACTCGGTGGTCAAAATGTTGTGCGCTGCCAGCTGGTTAATGCTTAATCTTCC 4826
Db 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY 4827 AGAATGCTCCAGGGAAGAGACTAAGGGAAGCCACCCAGGCGCTCACTGGGAAGTGAC 4886
Db 841 ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
QY 4887 TTCCTAGGTAAAGCCGCTAAATACGGAACAAATACCTATTGGTTTTCTAGACACC 4946
Db 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
QY 4947 TTTTCAGGATGGGTAGAGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTCTAAA 5006

Db 881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
QY 5007 AAAATACTCGAAGAAATTTTCCAGATTTGGGAATACCTAAAGTAAATAGGCTCAGACAAT 5066
Db 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
QY 5067 GGTCCAGCTTTTGTGTTCCCGCAGGTAACTCAGGGACTCGCCCAAGATATTGGGGATTGATTGG 5126
Db 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY 5127 AAATGTCATTGTGCATACAGACCCCAAGACTCAGGACAGTAGAGAGGATGAATAGAAC 5186
Db 941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY 5187 ATTAAGAGACACCTTACTAAATTTGACCGGAGACTGGCGTTAATCATTTGGATAGCTCTC 5246
Db 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
QY 5247 CTGCCCTTTGTGCTTTTAGGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTCAT 5306
Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
QY 5307 GAATTAATCTACGGGGACCCCTCCATTTGGTAGAAATTTGCTTCTCTACATAGTCTGAC 5366
Db 1001 GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
QY 5367 GTGCTGCTTTCCAGCGCTTTGTTCTCTAGCTCAAGGACTTGAAGGCTGAGACACGA 5426
Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY 5427 GGTGAGGAGCAACTCCGGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCACATCGT 5486
Db 1041 AlaTrpArgGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnIleProHisArg 1060
QY 5487 TTCCAAGTGGAGAGATTAGCTTACGTTAGACGCCCGCTGAGGAAACCTCGAGACTCGG 5546
Db 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
QY 5547 TCGAAGGGCCCTTATCTGCTACTTTTACCACACCAACCGCTGTGAAAGTCGAAGGAATC 5606
Db 1081 TrpLysGlyProTyrLeuValLeuLeuThrProThrAlaValLysValGluGlyIle 1100
QY 5607 TCCACTGTGATCCATCCACCGTTAAACCGGCGCCACCTCCCGATTCCGGGTGGA 5666
Db 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1120
QY 5667 GCCAAAAGACTGAAAATCCCTTAAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTC 5726
Db 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
QY 5727 AATACCTCTCAGAC 5741
Db 1141 AsnAsnLeuSerAsp 1145
RESULT 4
ADS73447
ID ADS73447 standard; protein; 1145 AA.
XX ADS73447;
DT 16-DEC-2004 (first entry)
XX Swine retroviral pol protein.
DE Swine retroviral pol protein.
XX Swine retroviral protein; immunosuppressive; gene therapy; pol protein.
KW Porcine endogenous retrovirus.
XX OS
XX US2004185435-A1.
XX PN
XX 23-SEP-2004.
XX PD
XX 26-NOV-2003; 2003US-00723552.

XX 14-DEC-1995; 95US-00572645.
 PR 13-DEC-1996; 96US-0076528.
 PR 14-SEP-2000; 2000US-00661858.
 XX (GEO) GEN HOSPITAL CORP.
 XX PA
 XX PI Fishman JA;
 XX WPI; 2004-689179/67.
 DR N-PSDB; ADS73369.
 XX
 PT New porcine retroviral polypeptide encoded by a nucleic acid, useful in
 PT evaluating an immunosuppressive treatment for the ability to activate a
 PT retrovirus, such as an endogenous porcine retrovirus.
 XX
 PS Claim 8; Fig 3; 83pp; English.
 XX
 CC The present invention relates to the swine retroviral polypeptides and
 CC their encoding nucleic acids. The methods and compositions of the present
 CC invention are useful for screening a cell or tissue, e.g. a heart, lung,
 CC liver, bone marrow, kidney, brain cells, neural tissue, pancreas and
 CC intestinal tissue xenograft, for the presence or expression of a swine or
 CC miniature swine retrovirus or retroviral sequence. The invention is also
 CC useful in evaluating an immunosuppressive treatment for the ability to
 CC activate a retrovirus such as an endogenous porcine retrovirus. The
 CC invention is also useful in gene therapy. The present sequence is the
 CC swine retroviral pol protein.
 XX
 SQ Sequence 1145 AA;
 XX
 Alignment Scores:
 Pred. No.: 0 Length: 1145
 Score: 6038.00 Matches: 1145
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 41.3% Indels: 0
 DB: 8 Gaps: 0
 US-10-723-552-3 (1-8132) x ADS73447 (1-1145)

QY	2787	TCCTCCCTGGAATACCTCCCTGCTACCGGTTAGAAAGCCCTGGGACTAATGACTATCGACCA	2846
DB	161	SerProTrpAsnThrProLeuLeuProValAlaGlyProGlyThrAsnAspTyrArgPro	180
QY	2847	GTACAGGACTTGAGAGAGGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAAC	2906
DB	181	ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn	200
QY	2907	CCTTAATACCTCTTGCTGCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGACTTA	2966
DB	201	ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu	220
QY	2967	AAGGATGCTCTCTCTGCTGAGATTACACCCCACTAGCAACCACTTTTTCGCTTCGAA	3026
DB	221	LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu	240
QY	3027	TGGAGAGATCCAGGTACGGGAAGAACCCGGCAGCTCACCTGGACCCGACTGCCCCAAAGG	3086
DB	241	TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly	260
QY	3087	TTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGACCTGGCCAACTTCAGG	3146
DB	261	PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg	280
QY	3147	ATCCAAACACCTCAGGTGACCCCTCCCTCCAGTACGTGATCAGCTGCTTCTGGCGGGAGCC	3206
DB	281	IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla	300
QY	3207	ACCAAAACAGGACTGCTTAGAAGGCACGAAGGCACTACTGCTGAAATGCTGACCTAGGC	3266
DB	301	ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAspLeuGly	320
QY	3267	TACAGAGCTCTGCTAAGAAGCCCGACAGATTGACGAGAGAGTAACTATCTTTGGGGTAC	3326
DB	321	TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr	340
QY	3327	AGTTTCGGGACGGCAGCGATGCGCTGACGAGGCACGGAAGAAAACCTGTAGTCCAGATA	3386
DB	341	SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle	360
QY	3387	CGGGCCCCAACCAACAGCCAAACAAATGAGAGAGTATTTTGGGGAACAGCTGGATTTGCGA	3446
DB	361	ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg	380
QY	3447	CTGTGATCCCGGGTTTGGGACCTTAGCAGCCCACTTACCGCTACCCAAAGAAAA	3506
DB	381	LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys	400
QY	3507	GGGGAATTCCTCGGCTCCTGAGCACCAAGGCAATTTGATGCTATCAAAAAGGCCCTG	3566
DB	401	GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu	420
QY	3567	CTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCCCTTTACCTTTATGTGGAT	3626
DB	421	LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp	440
QY	3627	GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCCAACCCCTAGAGCCATCGGAGAGACCT	3686
DB	441	GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro	460
QY	3687	GTGCGCTACTCTGTCAAAGAAAGCTCGATCCTGTAGCCAGTGGTTGGGCCCATATGCTGAAG	3746
DB	461	ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLys	480
QY	3747	GCTATCCGAGCTGTGGCCATACCTGGTCAAGGACGCTGACAAATAGCTTTGGGACAGAAAT	3806
DB	481	AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn	500
QY	3807	ATNACTGTAAATAGCCCCCATGCAATGGAGAACTGTTGGCAGCCCCCAGACCCGATGG	3866
DB	501	IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp	520
QY	3867	ATGACCAACGCCCGCATGACCCCACTATCAAAAGCCTGCTTCTTCACAGAGAGGTCACGTTTC	3926

```

Db      521 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540
QY      3927 GCTCCACGACGCGCTCTCAACCCCTGCCACTCTTCTGCTCAGAGACTGATGAACAGTG 3986
Db      541 AlaProAlaAlaLeuAsnProAlaThrLeuLeuLeuProGluThrAspGluProVal 560
QY      3987 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGTCCGCAAGGACCTTACAGAC 4046
Db      561 ThrHisAspCysHisGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 580
QY      4047 ATACCGCTGACTGCGAAGGCTTAACCTGTTCACTGACCGAAGCAGCTATGCTGGAA 4106
Db      581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
QY      4107 GGTAAGAGGATGGCTGGGCGCGGGTGTGCGAGCGGACCGCCAGCATCTGGCGCAGCAGC 4166
Db      601 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
QY      4167 CTGCGGAAGGAATCTTCAGCACAAAGGCTGAGCTCATGGCCCTCAGCGCAAGCTTTGCGG 4226
Db      621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY      4227 CTGCGCAAGGGAATCCATAACNTTATACGACAGCAGGATGCTTGGCTGCGACTGCA 4286
Db      641 LeuAlaGlnGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
QY      4287 CAGCTACATGGGCGCATCTATAACCAAGGGGTGTCTTAACCTCAGCAGGAGGGAATA 4346
Db      661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY      4347 AAGAACAAAGAGAAATTTAAGCCCTATTAGAAGCCGTACATTTACCAAAAGGCTAGCT 4406
Db      681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
QY      4407 ATTATACACTGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATG 4466
Db      701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
QY      4467 GCTGACCGGGTTCCTCAAGCAGGCGCCCGGGTGTAAACCTCTGCTTAAATAGAAATG 4526
Db      721 AlaAspArgValAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProIleIleGluMet 740
QY      4527 CCAAAAGCCCAAGACCCAGACAGACAGTACACCTAGAAAGACTGGCAAGAGATAAAAAAG 4586
Db      741 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
QY      4587 ATAGACAGTTCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGGAATC 4646
Db      761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
QY      4647 CTGCCCCACAAGAAGGGTTAGAATATGTCCAACAGATACATCGTTAAACCCACCTAGGA 4706
Db      781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY      4707 ACTAAACACTGCGAGTGTGTGAGAACATCCCTTATCATGTCTTGAGGCTTACCAGGA 4766
Db      801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
QY      4767 GTGGCTGACTCGGTGTCAACATTTGTGTGCTGCGCTGCGACCTGGTTAATCTAATCTTCC 4826
Db      821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY      4827 AGAATCCCTCCAGGGAAGAGACTAAGGGAAGCCACCGAGGCGCTCAGTGGGAAGTGGAC 4886
Db      841 ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
QY      4887 TTCCTAGAGTAAGCCGCTAAATACGGAACAAATACCTATTGTTTTGTAGACACC 4946
Db      861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
QY      4947 TTTTCAGGATGGTAGAGCTTATCTCTACTAGAAAGAGACTTCAACCGTGTGGCTAAA 5006

```

```

Db      881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
QY      5007 AAAATACTGGAAGAAAATTTTTTCCAGATTTGGAAATACCTAAGCTAATAGGTCAGACAAT 5066
Db      901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
QY      5067 GGTCCAGCTTTTGTTCCTCCAGGTAACTCAGGGACTGGCCCAAGATATTGGGGATTGATTGG 5126
Db      921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY      5127 AAATCTCATGTGCTACAGACCCCAAGACTCAGGACAGGTAGAGAGATGAATAGAAC 5186
Db      941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY      5187 ATTAAGAGACCTTACTAAATTTGACCGCGAGACTGGGGTAAATGATTGGATAGCTCTC 5246
Db      961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
QY      5247 CTGCCCTTTGTGCTTTTGTAGGGTTAGGAAACACCCCTGGACAGTTTGGGCTGACCCCTAT 5306
Db      981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
QY      5307 GAATTACTCTACGGGGGACCCCTCCCTTGGTAGAAATTTGCTTCTGTACTAGTCTGCTGAC 5366
Db      1001 GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
QY      5367 GTGCTCTTTTCCAGGCTTTGTTCTTAGGCTCAAGGCACTTGAGTGGGTGAGACAACGA 5426
Db      1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY      5427 GCGTGGAGGCAACTCGGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCACATCGT 5486
Db      1041 AlaTrpArgGlnLeuArgGluAlaTyrSerGlyGlyLysLeuGlnIleProHisArg 1060
QY      5487 TTCCAAGTGGAGATTCACTAGTTAGACGCCACCGTCGAGGAAACCTCGAGACTCGG 5546
Db      1061 PheGlnValGlyAspSerValTyrValArgHisArgAlaGlyAsnLeuGluThrArg 1080
QY      5547 TGGAGGGCCCTTATCTCGTACTTTTACCACACCAACCGCTGTGAAAGTCGAAGGAATC 5606
Db      1081 TrpLysGlyProTyrLeuValLeuLeuThrProThrAlaValLysValGluGlyIle 1100
QY      5607 TCCACTGGATCCATCCATCCCACTTAAACCGGCGCCACCTCCCGATTTCGGGGTGGAAA 5666
Db      1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1120
QY      5667 GCGCAAAAGACTGAAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCTTACTCTGTC 5726
Db      1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
QY      5727 AATAACCTCTCAGAC 5741
Db      1141 AsnAsnLeuSerAsp 1145

```

RESULT 5

```

AAB10044
ID AAB10044 standard; protein; 1737 AA.
XX
AC AAB10044;
XX
DT 02-NOV-2000 (first entry)
XX
DE MMLV gag-pol protein.
XX
KW Glycoprotein; gag gene; pol gene; GP-1; GP-2; anti-HIV; cytostatic;
KW gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;
KW carcinoma; melanoma.
XX
OS Moloney murine leukemia virus.
XX
PN EP1006196-A2.
XX
PD 07-JUN-2000.

```

XX 25-NOV-1999; 99EP-00250415.
 XX 26-NOV-1998; 98DE-01056463.
 PR (PETT-) PETTE INST HEINRICH.
 XX Von Laer MD;
 XX WPI; 2000-378268/33.
 DR New retroviral packing cell useful as pharmaceutical carrier in gene
 PT therapy for treatment of HIV and neoplasms, comprises retroviral genes
 PT and glycoproteins.
 XX Disclosure; Page 32-36; 69pp; German.
 XX This invention describes a novel retroviral packing cell (I), comprising
 CC the retroviral genes gag, pol and glycoproteins gp-1 and gp-2 of the LCMV
 CC coding gene gp, or a part of these. The products of the invention have
 CC anti-HIV and cytostatic activity and can be used for gene therapy. (I) is
 CC useful for in vitro infection of cells, especially hematopoietic stem
 CC cells, for expression of transgenes in cells and as a pharmaceutical
 CC carrier for gene therapy. (I) is therefore useful in the treatment of
 CC infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and
 CC other diseases. This sequence represents the Moloney murine leukemia
 CC virus (MMLV) gag-pol protein which is described in the method of the
 CC invention
 XX Sequence 1737 AA;
 SQ

Alignment Scores:
 Pred. No.: 0 Length: 1737
 Score: 5338.00 Matches: 1035
 Percent Similarity: 71.6% Conservative: 239
 Best Local Similarity: 58.2% Mismatches: 395
 Query Match: 36.5% Indels: 110
 DB: 3 Gaps: 26

US-10-723-552-3 (1-8132) x AAB10044 (1-1737)

QY 585 ATGGGACAGCGGTGACGACCCCTCTTAGTTGACTCTGACCATCGACTGCAAGTAA 644
 DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
 QY 645 TCAGGGCTCATATTTGTCAGTTAGGTTAAGAGGACCTTGGCAGACTTCTGTGTC 704
 DB 21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgTrpValThrPheCysSer 40
 QY 705 TCTGATGGCCGACATTTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATCTGAGATT 764
 DB 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
 QY 765 ATCCTGGCTCTTAAACAGATTATTTTTCAGACTGGACCCCGCTCTCATCCCGATCAGGAG 824
 DB 61 IleThrGlnValLysLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
 QY 825 CCTATATCTTACGTGGCAAGATTGGCGAGAGATCTCCGCCATCGGTTAAACCATGG 884
 DB 81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProTrpValLysProPhe 100
 QY 885 CTGAATAAGCCCAAGAACCCAGGTCCCGAATTCCTGCTCTTGGAGAGAAACCAACAC 944
 DB 101 ValHis---ProLysProProProProLeuProProSerAlaProSerLeuProLeuGlu 119
 QY 945 TCGGCTGAATAAGTCACGCCCTCTCTCATATCTACCCCGAGATTGAGGACCA---CCG 1001
 DB 120 ProProArgSerThrProProArgSerSerLeuTyrProAlaLeuThrProSerLeuGly 139
 QY 1002 GCTTGGCCGGAACCCCA-----TCT 1022
 DB 140 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuIleAspLeuLeuThr 159

QY 1023 GTTCCCCCACCCTTATCTGCGACAGGGTCCGCGAGGGAGCCCTTTGCCCTCTCT--- 1079
 DB 160 GluAspProProTyr-----ArgAspProArgProProSer 173
 QY 1080 -----GGAGCTCGGCGGTGGAGGACCT----- 1103
 DB 174 AspArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyGluAlaProAspProSerPro 193
 QY 1104 ---GCTGAGGACTCGGAGCGGGGGCGCCACCCTCGGAGCGGACAGAGATCGCG 1160
 DB 194 MetAlaSerArgLeuArgGlyArgGluProProValAlaAspSerThrThrSerGln 213
 QY 1161 ACATTACCCTGCGCACGTACGCGCTCCACACCGGGGGGCAATGCGACCCCTCCAG 1220
 DB 214 AlaPheProLeuArgAlaGlyGly-----AsnGlyGln-----LeuGln 226
 QY 1221 TATTGGCCCTTTCTCTCAGATCTCTATATATTGAAAACTAACCATCCCTTTCTCG 1280
 DB 227 TyrTrpProPheSerSerAspLeuTyrAsnTrpLysAsnAsnProSerPheSer 246
 QY 1281 GAGGATCCCAACGCTCAGCGGTGGTGGAGTCCCTTATGTTCTCTCACAGCCTACT 1340
 DB 247 GluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuThrHisGlnProThr 266
 QY 1341 TGGGATGATTGCAACAGCTGCTGCAGACACTCTTCCACACCGAGGAGCGAGAGAATT 1400
 DB 267 TrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGlyGlnArgVal 286
 QY 1401 CTATTAGAGGCTAGAAAAAATGTTCTGCGGCGCGAGCGGCGACCCACGCGGTGCAAAAT 1460
 DB 287 LeuLeuGluAlaArgLysAlaValArgLysAspAspGlyArgProThrGlnLeuProAsn 306
 QY 1461 GAGATTGACATGGATTTCCTTACTCGCCCGGTTGGGACTTACAAACGCGTGAAGGT 1520
 DB 307 GluValAspAlaAlaProLeuGluArgProAspTrpAspTrpThrGlnAlaGly 326
 QY 1521 AGGGAGAGCTTGAAAAATCTATCGCCAGGCTCTGGTGGCGGGTCTCGGGCGCGCTCAAGA 1580
 DB 327 ArgAsnHisLeuValHisTyrArgGlnLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 346
 QY 1581 CGGCCCACTAATTTGGCTAAGTAAGTAAGTATGTCAGGAGCCGAATGAACCCCTCT 1640
 DB 347 SerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer 366
 QY 1641 GTTTTCTTGGAGGCTCTTGAAGCCTTCAGCGGTACACCCCTTTTGTATCCACCTCA 1700
 DB 367 AlaPheLeuGluArgLeuLysGluAlaTyrArgTyrThrProTyrAspProGluAsp 386
 QY 1701 GAGGCCCAAAAGCCCTCAGTGGCTTTGGCCTTTATAGGACAGTCAGCTTGGATATTAGA 1760
 DB 387 ProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly 406
 QY 1761 AGAAGCTTCAGAGCTGGAAGGTTACAGGAGGCTGAGTTAGCTGATCTAGTAGAGGAG 1820
 DB 407 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 426
 QY 1821 GCAGAGAACTATATTACAAAAGGAGACAGAGAGAAAGGCAACAAAGAAAAGAGA 1880
 DB 427 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgGly 446
 QY 1881 GAAAGAGAGAAAGGAGGAGAAAGA-----CGT 1907
 DB 447 GluThrGluGluLysGluGluArgArgThrGluAspGluGlnLysGluLysGluArg 466
 QY 1908 AATAAACGCGACAGAGAAGAAATTGACTAAGATCTTGGCTGCAGTGGTTGAGGAAAAAGC 1967
 DB 467 AspArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys 486
 QY 1968 AATACGGAAGAGAGAGAGATTTTAGGAAAATTTAGTTCAGGCCCTAGACAGTCAGGGAAC 2027
 DB 487 Gln-----AspArgGlnGlyGlyGlu 493
 QY 2028 CTGGGCAATAGGACCCCTCTCGCAAGGACCAATGTGCATATTGTAAAGAAAGAGGACAC 2087

Db 494 ---ArgArgSerGlnLeuAspArgAspGlnCysAlaTyrCysLysGluLysGlyHis 512
QY TGGCCAAAGGAACCTGCCCAAGAG-----GGAACAACAAGGACCAAGG----- 2129
Db 513 TrpAlaLysAspCysProLysLysProArgGlyProArgGlyProGlnThrSer 532
QY ATCTAGCTCTAGAGAGATAAAGATTAGGGAGACGGGGTTCGACCCCTCCCGAG 2189
Db 533 LeuLeuThrLeuAspAsp-----GlyGlyGlnGlnGluProProGlu 548
QY CCCAGGTAACCTTTGAAGTGGAGGCGACCACTGAGTTCCTGCTTGTATACCGGCG 2249
Db 549 ProArgIleThrLeuLysValGlyGlnProValThrPheLeuValAspThrGlyAla 568
QY AAACATTTCAGTGTACTACAGCCATTAGGAAAAATAAAGATAAAAAATCTGGGTGATG 2309
Db 569 GlnHisSerValLeuThrGlnAsnProGlyProLeuSerAspLysSerAlaTrpValGln 588
QY GTGCGCACAGGGCAACAACAGTATCCATGACTACCGAAGAACAGTTCGATTTGGAGTG 2369
Db 589 GlyAlaThrGlyGlyLysArgTyrArgTrpThrAspArgLysValHisLeuAlaThr 608
QY GGAGGGTAACCACTCGTTCTGCTCATCTGAGTGGCCAGCACCCCTCTTAGTGA 2429
Db 609 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 628
QY GACTTATTGACCAAGATGGGACACAAATTTCTTTTGA-----CAAGGGAACACAGAAGTG 2486
Db 629 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValMet 648
QY TCTGCAAAATAACAACCTATCACTGTGTGACCTCCAATTAGATACGAATATGACTA 2546
Db 649 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu 668
QY TACTCTCCCTAGTAAGCTGTATCAAAATATA---CAATTCTGGTTGGAACAGTTTCCC 2603
Db 669 HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 688
QY CAAGCTCGGCGAACAACCGAGGATGGGTTTGGCAAAAGCAAGTTCCTCCCAAGTTATT 2663
Db 689 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 708
QY CAACTGAAGCGCAGTCCACACAGTGTCACTGACAGACAGTACCCCTTGAGTAAAGACT 2723
Db 709 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrProMetSerGlnGluAla 728
QY CAAGAAGGAATTCGGCGCATGTCGAAAGATTATCAACAGGGCATCTAGTTCTGTC 2783
Db 729 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 748
QY CAATCTCCCTGGAATACTCCCTGCTACCGGTTAGAAAGCCCTGGGACTAATCACTATCGA 2843
Db 749 GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 768
QY CCAGTACAGGACTTGAGAGAGTCAATAACCGGGTCGAGGATATACACCAACAGTCCCG 2903
Db 769 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 788
QY AACCTTTAATACCTCTTGTGTGCTCTCCACCCCAACGAGGCTGGTATACAGTATTGGAC 2963
Db 789 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp 808
QY TTAAAGGATGCTTCTTCTGCTGAGATTACCCCTAGCAACCACTTTTTCCTTC 3023
Db 809 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 828
QY GAATGGAGAGATCCAGGTACGGGAAGAACCGGCGCAGCTCACCTGGACCCGACTGCCCAA 3083
Db 829 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 848
QY GGGTTCAAGAACTCCCGGACCATCTTTGACGAAGCCCTTACAGAGACCTGGCCAACTTC 3143

Db 849 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 868
QY AGGATCCAAACACACCTCAGGTGACCTCTCCAGTACGTGATGATCACTGCTTCTGGCGGA 3203
Db 869 ArgIleGlnHisProAspLeuIleLeuLeuGlnTyrValAspAspLeuLeuAlaAa 888
QY GCCACCAACAGAGCTGCTTAGAAGCACGAAGGCACTACTGCTGGAATTTGCTGACCTA 3263
Db 889 ThrSerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu 908
QY GCTACAGAGCTCTGCTTAAGAGGCCAGCTTTCAGGAGAGAGAGCTTAACATCTTTGGG 3323
Db 909 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 928
QY TACAGTTTCGGGACCGGCGAGCGATGGCTGACGAGGACGAGGACGGAAGAACTGTAGTCCAG 3383
Db 929 TyrLeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGlnThrValMetGly 948
QY ATACCGGGCCCCAACACACGCCAAACAAATGAGAGAGTTTTTGGGGACAGCTGGATTTGC 3443
Db 949 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 968
QY AGACTGTGATCCCGGGTTTCGGACCTTAGCAGCCCCACTCTACCGCTAACCAAGAA 3503
Db 969 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 988
QY AAAGGGGAATCTCTCGGCTCTGAGCACACAGAGGAGCATTTGATCTATCAAAAAGGCC 3563
Db 989 GlyThrLeuPheAsnTrpGlyProAspGlnGlnLysAlaTyrGlnGluIleLysGlnAla 1008
QY CTGCTGAGCGCACTGCTCTGCCCCCTCTGAGCTAACTAAACCTTTTACCTTTTATGTG 3623
Db 1009 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 1028
QY GATGAGCGTAAGGGAGTAGCCCGGGAGTTTTAAACCAACCTCTAGGACATGGAGAA 3683
Db 1029 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArg 1048
QY CTGCTCGCTACTCTCAAGAGCTCGATCTCTGAGCAGTGGTTGGCCCATATGCTG 3743
Db 1049 ProValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTrpProCysLeu 1068
QY AAGGCTATCGCAGCTGTGTGCCATCTGCTCAAGGACGCTGACAAATTTGATTTGGACAG 3803
Db 1069 ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln 1088
QY AATATACTGTATAGCCCCCATGTCATTTGGAGAACATCGTTCCGGCAGCCCCCAGACCGA 3863
Db 1089 ProLeuValIleLeuAlaProHisAlaValGluAlaLeuValLysGlnProAspArg 1108
QY TGGATGACCAAGCCCGCATGACCCACTATCAAGCCTGCTTCTC---ACAGAGAGGCTC 3920
Db 1109 TrpLeuSerAsnAlaArgMetThrHisTyrGlnAlaLeuLeuLeuAspThrAspArgVal 1128
QY AGTTCTGCTCCACAGCCGCTCTCAACCTCGCACTCTTCTGCTGGAAGAGACTGATGAA 3980
Db 1129 GlnPheGlyProValValAlaLeuAsnProAlaThrLeuLeuPro---LeuProGluGlu 1147
QY CAGTGTACTCATGATGCTCATCTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTT 4040
Db 1148 GlyLeuGlnHisAsnCysLeuAspIleLeuAlaGluAlaHisGlyThrArgProAspLeu 1167
QY ACAGACATACCGCTCACTGGAGAGTGTCAACCTGCTTCTGCTGAGGAGAGCTATGTC 4100
Db 1168 ThrAspGlnProLeuProAspAlaAspHisThrTrpTyrThrAspLysSerSerLeuLeu 1187
QY GTGAAGTAAGAGATGCTGGGCGCGGTGGTGGAGCGGACCCGACGATCTGGGCC 4160
Db 1188 GlnGluGlyGlnArgLysAlaGlyAlaValThrThrGluThrGluValIleTrpAla 1207
QY AGCAGCTTCGGGAAGGAACTTCAGCACAAAAGGCTGAGCTCATGGCCCTCAGCAGCT 4220
Db 1208 LysAlaLeuProAlaGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAla 1227

QY 4221 TTCCGGCTGCGCAGGAATCCATTAACGACAGCAGGTATGCTTTGGG 4280
 Db 1228 LeuLysMetAlaGluGlyLysLysLeuAsnValThrAspSerArgTyrAlaPheAla 1247
 QY 4281 ACTGCACACGTACATGGGGCCATCTATAACAAGGGGTTGCTTACCTCAGCAGGAGG 4340
 Db 1248 ThrAlaHisIleHisGlyLysLysLysArgArgGlyLeuLeuThrSerGluGlyLys 1267
 QY 4341 GAAATAAAGAACAAAGAGGAATCTAAGCCTATTAGAGCCCTACATTATCCAAAAGG 4400
 Db 1268 GluIleLysAsnLysAspGluIleLeuAlaLeuLysAlaLeuPheLeuProLysArg 1287
 QY 4401 CTAGCTATTATACACTGCTGACATCAGAAAGCTAAGATCTCATATCCAGAGAAC 4460
 Db 1288 LeuSerIleIleHisCysProGlyHisGlnLysGlyHisSerAlaGluAlaArgGlyAsn 1307
 QY 4461 CAGATGGCTGACCGGGTTGCCAAGCAGGAGCCAGGGTGTAAACCTTCTGCCTATAATA 4520
 Db 1308 ArgMetAlaAspGlnAlaAlaArgLysAlaAla-----IleThr 1320
 QY 4521 GAAATGCCCAAGACC----- 4547
 Db 1321 GluThrProAspThrSerThrLeuLysGluAsnSerSerProTyrThrSerGluHis 1340
 QY 4548 CGACAGTACACCTAGACACTGGCAGAGATAAAGATAGACACAG--TTCTCTGAG 4604
 Db 1341 PheHisTyrThrValThrAspIleLysAspLeuThrLysLeuGlyAlaIleTyrAspLys 1360
 QY 4605 ACTCCGGAAGGGACCTGCTATACCTCAGATGGGAAGAAATCTGCCCCACAAAGAGGG 4664
 Db 1361 ThrLysLysTyrTrpValTyr-----GlnGlyLysProValMetProAspGlnPheThr 1378
 QY 4665 TTAGAATATGTCACAGATACATCGTCTAAACCCACTAGGAATAACACCTGCAGCAG 4724
 Db 1379 PheGluLeuLeuAspPheLeuHisGlnLeuThrHisLeuSerPheSerLysMetLysAla 1398
 QY 4725 TTGGTCAGACA-----TCCCTTATCATGTTCTGAGGCTACAGAGTGCGTGACTCG 4778
 Db 1399 LeuLeuGluArgSerHisSerProTyrTyrMetLeuAsnArgAspArgThrLeuLysAsn 1418
 QY 4779 GTGGTCAAAACATGTGTGCCCTGCCAGCTGTTAAATGCTAATCTTCCAGAAATGCCTCCA 4838
 Db 1419 IleThrGluThrCysLysAlaCysAlaGlnValAsnAlaSerLysSerAlaValLysGln 1438
 QY 4839 GGAAGAGACTAAGGGAAGCCACCCAGGCGCTCACTGGGAATGGACTTCACTGAGGTA 4898
 Db 1439 GlyThrArgValArgGlyHisArgProGlyThrHisTyrGluIleAspPheThrGluIle 1458
 QY 4899 AAGCCGGCTAAATACGMAACAATACCTATTGTTTGTAGACACCTTTTCAGAGTG 4958
 Db 1459 LysProGlyLeuTyrGlyTyrLysTyrLeuLeuValPheIleAspThrPheSerGlyTrp 1478
 QY 4959 GTAGAGCTTATCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAAAAATACTCGAA 5018
 Db 1479 IleGluAlaPheProThrLysLysGluThrAlaLysValValThrLysLysLeuLeuGlu 1498
 QY 5019 GAAATTTTCCAAAGATTGGAATACCTAAGTAAATAGGTCAGACAATGTCAGCTTTT 5078
 Db 1499 GluIlePheProArgPheGlyMetProGlnValLeuGlyThrAspAsnGlyProAlaPhe 1518
 QY 5079 GTTGGCCAGTAACTCAGGACTGGCCAGATATTGGGATTCATTGGAACTGCAATCTT 5138
 Db 1519 ValSerLysValSerGlnThrAlaAspLeuLeuGlyIleAspTrpLysLeuHisCys 1538
 QY 5139 GCATACAGACCCCAAGCTCAGACAGGTAGAGAGGATGAATAGAACCATTAAGAGACC 5198
 Db 1539 AlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrLysLysGluThr 1558
 QY 5199 CTTACTAAATGACCGCGAGACTGGCGTTAATGATTGATGACTCTCTGCGCCTTTG 5258
 Db 1559 LeuThrLysLeuThrLeuAlaThrGlySerArgAspTrpValLeuLeuLeuProLeuAla 1578

QY 5259 CTTTTTAGGTTAGGAACACCCCTGGCAGTGTGGCTGACCCCTATGAATTACTCTAC 5318
 Db 1579 LeuTyrArgAlaArgAsnThrProGlyProHisGlyLeuThrProTyrGluIleLeuTyr 1598
 QY 5319 GGGGACACCCCCCATTTGGTAGAAATTCCTTCTGTACATAGTCTGACGTGCTTTCC 5378
 Db 1599 GlyAlaProProProLeuValAsnPheProAspProAspMetThrArgValThrAsnSer 1618
 QY 5379 CAGCCTTTGTTCTTAGGCTCAAGGCTTCAAGGCTTGTAGTGGTGAGACAACGAGGTGGAGCAA 5438
 Db 1619 ProSerLeuGlnAlaHisLeuGlnAlaLeuTyrLeuValGlnHisGluValTrpArgPro 1638
 QY 5439 CTCGGGAGGCGCTACTCAGGAGGAGAGACTTGCAG--ATCCCATCGTTTCCCAAGTG 5495
 Db 1639 LeuAlaAlaAlaTyrGlnGlnLeuAspArgProValValProHisProTyrArgVal 1658
 QY 5496 GGAGATTACGTCTACGTTAGACCCACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGCG 5555
 Db 1659 GlyAspThrValTrpValArgHisGlnThrLysAsnLeuGluProArgTrpLysGly 1678
 QY 5556 CTTATCTCTGACTTTTGACACCAACCGCTGTGAAGTCAAGGAATCTCCACCTGG 5615
 Db 1679 ProTyrThrValLeuLeuThrThrProThrAlaLeuLysValAspGlyIleAlaAlaTrp 1698
 QY 5616 ATCCATGCTATCCACGTTAAACCGCGCCACCT-----CCCGATTGCG-----GGG 5660
 Db 1699 IleHisAlaAlaHisValLysAlaAlaAspProGlyGlyProSerSerArgLeuThr 1718
 QY 5661 TGAAGAAGCCGAAAGACTGAAATCCCTTAAAGTTCGCTTCATCGCTGGTTCCT 5717
 Db 1719 TrpArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGluAlaPro 1737
 RESULT 6
 AAR94427
 ID AAR94427 standard; protein; 1784 AA.
 XX AAR94427;
 XX 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 11-JUN-1996 (first entry)
 XX FeLV F6A provirus clone 61E encoded nucleocapsid protein.
 XX FeLV; retrovirus; vaccine; AIDS; disease model; immunodeficiency;
 KW viraemia; leukaemia; therapy; nucleocapsid.
 XX Feline immunodeficiency virus; subtype A.
 XX Key Location/Qualifiers
 FT Misc-difference 581 /notes "codon 581 in encoding sequence is TAG"
 XX EP699758-A1.
 XX 06-MAR-1996.
 PF 12-DEC-1989; 95EP-00100871.
 XX 13-DEC-1988; 88US-00284139.
 XX (HARD) UNIV HARVARD.
 PA (COLS) UNIV COLORADO STATE RES FOUND.
 XX Hoover EA, Mullins JI;
 XX WPI; 1996-180826/19.
 DR N-PSDB; AAT13265.
 XX Inactivated FeLV-A sub-type isolates for use in disease models and
 PT vaccines - can be used to study prophylaxis and therapy of related immuno
 PT -deficiencies in other species, e.g. human.
 XX

	Qy	512	TCACGTGTCCTGGATCTGTTTCTGTTTCTGTTTGTGTCCTTGTCCTGTCGTCTGTCTGT	571
	Db	47	SerArgAlaAIAArgLeuValIlePheCysLeuValAlaSerPheLeuVal---ProCys	65
	Qy	572	CTACAGTTTT-----AAATGGGCAGACGGTGACGACCCCTCTAGTTT	616
	Db	66	Leu-ThrPheLeuIleAlaGluThrValMetGlyGlnThrIleThrThrProLeuSerIle	85
	Qy	617	GACTCTCGACCATTGGACTGAAGTTAAAATCCAGGGCTCATATAATTTGTCAAGTTACGTTAA	676
	Db	85	uThrLeuAspHisTrpSerGluValArgAlaargAlaHisasnGlnGlyValGluValar	105
	Qy	677	GAAGGGACCTTTGGCAGACCTTCTGTGTCCTGAATGGCCGACATTCGATGTTGGATGCC	736
	Db	105	glyLysIlysTrpIleThrLeuCysGluAlaGluTrpValMetMetAsnValGlyTrpPr	125
	Qy	737	ATCAGAGGGGACCTTAATTTCTGAGATTATCTCGCTCTTAAAGCAGGTATTTTTCAGAC	796
	Db	125	oArgGluGlyThrPheSerLeuAspAsnIleSerGlnValGluLysIlysIlePheAlaPr	145
	Qy	797	TGACACCGGCTCTCATCCCGATCAGGAGCCCTATATCTTACGTGCGCAAGTATTCGCAGA	856
	Db	145	oGlyProTyGlyHisProaspGlnValProTyIleThrThrTrpArgSerLeuAlath	165
	Qy	857	GGATCTCCGGCCATGGGTTAAACCATCGCTGAATAAGCCAAGAACCCAGGTCCCCGAAT	916
	Db	165	rAspProProSertRpValArgProPheLeuProProProIysThrProThrPro----	183
	Qy	917	TCTGGCTCTTGGAGNAAAAACAACACTCGGCTGAAAAGATCAAGCCCTCTCCT-----	971
	Db	184	-----LeuProGlnProLeuSerProGlnProSerAl	194
	Qy	972	-----CATATCTACCCGAGATT-----GAGGAGCCACCCGCG	1003
	Db	194	aProLeuThrSerSerLeuTyRProValLeuProLysSerAspProProLysProProva	214
	Qy	1004	TTGGCCGGAACCCCAATCTGTCCC-----CACACCCCTTA	1033
	Db	214	LeuProProAspProSerSerProLeuIleAspLeuLeuThrGluGluProProTy	234
	Qy	1040	TCTGGCACAGGTCCGGGAGGGAGACCTTTTGCCCTCTCGAGTCCG-----GCGGT	1093
	Db	234	r-----ProGlyGlyHisGlyPro-----ProProSerGlyProArGThrProTh	249
	Qy	1094	GGAGGACCTCTCAGGACTCGGAGCCGGAGGGGCCACCCCGGAGCGACAGACGA	1153
	Db	249	rAlaSerProIleAlaSerArgLeuArgGluArgGluAsnPro-----AlaGluGl	267
	Qy	1154	GATCGGACATTACCGCTGCGCAGCTACGGCCCTCCACACCGGGGGGCCAATTCGACC	1213

Db 602 oValThrPheLeuValAspThrGlyAlaGlnHisSerValLeuThrArgProAspGlyPr 622
QY 2282 ACTAAAGATAAAATCTCGGTGATGGTGCACAGGCCAACACAGTATCATGGAC 2341
Db 622 oLeuSerAspArgThrAlaLeuValGlnGlyAlaThrGlySerLysAsnTyrArgTrpTh 642
QY 2342 TACCCGAAGAACAGTTGACTGTGGAGTGGGACGGGTAAACCACTCGTTTCTGTGTCATACC 2401
Db 642 rThrAspArgValGlnLeuAlaThrGlyLysValThrHisSerPheLeuTyrValPr 662
QY 2402 TGAGTGCACAGCACCCCTCTTAGTGAGAGACTTATTGACCAAGATGGGACACAAATTC 2461
Db 662 oGluCysProTyrProLeuLeuGlyArgAspLeuLeuThrLysLysAlaGlnIleHi 682
QY 2462 TTTTGA---CAAGGGAACACAGAGTGTCTGCAATAACAACCTATCACTGTGTGAC 2518
Db 682 sPheThrGlyGluGlyAlaAsnValValGlyProArgGlyLeuProLeuGlnValLeuTh 702
QY 2519 CTTCCAAATTAGATGACGAATATCGACTATATCTCTCCCTAGTAAAGCCTGATCAAAATAT 2578
Db 702 rLeuGlnLeuGluGluGluTyrArgLeuPheGluProGluSerThrGlnLysGlnGluMe 722
QY 2579 ACAATTCTGTTGGAACAGTTTCCCAAGCCTCGGCAGAAACCGCAGGATGGTTTGGC 2638
Db 722 tAspIleTrpLeuLysAsnPheProGlnAlaTrpAlaGluThrGlyGlyMetGlyMetAl 742
QY 2639 AAGCAAGTTCCCCCAAGATTATTCACCTGAAGGCCAGTGCACACCACTGTCAGTCAG 2698
Db 742 aHisCysGlnAlaProValLeuIleGlnLeuLysAlaThrAlaThrProIleSerIleAr 762
QY 2699 ACAGTACCCCTCAGTAAAGAACTCAAGAGGAATTCGGCCCGCATGTCCAAAGATTAT 2758
Db 762 gGlnTyrProMetCProHisGluAlaTyrGlnGlyIleLysProHisIleArgMetLe 782
QY 2759 CCAACAGGGCATCTAGTCTCTGTCCATCTCCCTGGAATPACTCCCTGTCTACCGTTAG 2818
Db 782 uAspGlnGlyIleLeuLysProCysGlnSerProTrpAsnThrProLeuLeuProVally 802
QY 2819 AAGCCTGGGACTAATGACTATCGACCACTACAGGACTTGAGAGAGTCAATAAACGGGT 2878
Db 802 sLysProGlyThrLysAspTyrArgProValGlnAspLeuArgGluValAsnLysArgVa 822
QY 2879 GCAGGATATACACCAAGCTCCGAACCTTATACTCTGTGTCTCTCCACCCCA 2938
Db 822 lGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerThrLeuProProSe 842
QY 2939 ACGGAGCTGGTATACAGTATTGAGCTTAAAGTAGCCTCTCTCTGCTGAGATTACACCC 2998
Db 842 rHisProTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisSe 862
QY 2999 CACTAGCCAAACCACTTTTTCCTTCGATGGAGAGATCCAGGTACCGGAAGAACCGGCA 3058
Db 862 rGluSerGlnLeuLeuPheAlaPheGluTrpArgAspProGluIleGlyLeuSerGlyGl 882
QY 3059 GCTCACTGGACCCGCTGCCCAAGGTTCAAGACTCCCGCCACCATCTTTTCACGAAGC 3118
Db 882 nLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrThrPheAspGluAl 902
QY 3119 CTTACACAGAGACTGCGCAACTTCAGGATCCACACACCTTCAGGTGACCTCTCCAGTA 3178
Db 902 aLeuHisSerAspLeuAlaAspPheArgValArgTyrProAlaLeuValLeuLeuGlnTy 922
QY 3179 CGTGGATGACTGCTTCTGGCGGAGCCACCAACAGGACTGCTTAGAGGCGACGAAGGC 3238
Db 922 rValAspAspLeuLeuAlaAlaThrArgThrGluCysLeuGluGlyThrLysAl 942
QY 3239 ACTACTGCTGGAATTGCTCACCTAGCTACAGGCTCTGCTAAGAGGCCACAGATTTC 3298
Db 942 aLeuLeuGluThrLeuGlyAsnLysGlyTyrArgAlaSerAlaLysLysAlaGlnIleCy 962
QY 3299 CAGGAGAGGTAACACTTGGGGTACAGTTTTCGGGACGGGACGAGTGGCTGACGGA 3358
Db 962 sLeuGlnGluValThrTyrLeuGlyTyrSerLeuLysAspGlyGlnArgTrpLeuThrly 982

QY 3359 GGCACGGAAGAAAACTGTAGTCCAGATACCGGCCCCCAACCCACAGCCAAACAAATGAGAGA 3418
Db 982 sAlaArGlyGlyGluAlaIleLeuSerIleProValProLysAsnProArgGlnValArgGl 1002
QY 3419 GTTTTGGGGACAGCTGGATTTTTCAGACTGTGGATCCCGGGGTTTTCGCACTTAGCAGC 3478
Db 1002 uPheLeuGlyThrAlaGlyTyrCysArgLeuTrpIleProGlyPheAlaGluLeuAlaAl 1022
QY 3479 CCCACTCTACCGCTAACCAAGAAAAAGGGGAATTTCTCTGGGCTCTGTGAGCACCAGAA 3538
Db 1022 aProLeuTyrProLeuThrArgProGlyThrLeuPheGlnTrpGlyThrGluGlnGlnLe 1042
QY 3539 GGCATTGTGATGCTATCAAAAAGCCCTGCTGAGCGCACCTGTCTCTGCGCCCTCCCTGACGT 3598
Db 1042 uAlaPheGluAsnIleArgLysAlaLeuLeuSerSerProAlaLeuGlyLeuProAspI 1062
QY 3599 AACTAAACCTTTTACCCTTTATGTGATGAGCGTAAAGGAGTAGCCCGGGAGCTTTTAAC 3658
Db 1062 eThrLysProPheGluLeuPheIleAspGluAsnSerGlyPheAlaLysGlyValLeuVa 1082
QY 3659 CCAAAACCTTAGGACCATGAGAGAAGCTGTGCGCTACCTGTCAAAGAAGCTCGATCCTGT 3718
Db 1082 lGlnLysLeuGlyProTrpLysArgProValAlaTyrLeuSerLysLysLeuAspThrVa 1102
QY 3719 AGCCAGTGTTCGCCCATATGCTGAAGCTATCGCAGCTGTGGCCATCTACTGCTCAAGGA 3778
Db 1102 lAlaSerGlyTrpProCysLeuArgMetValAlaIleAlaIleLeuValLysAs 1122
QY 3779 CGCTGACAAATGACTTTGGGACAGAAATATACTGTAATAGCCCCCATCATGCTGGAGAA 3838
Db 1122 pAlaGlyLysLeuThrLeuGlyGlnProLeuThrIleLeuThrSerHisProValGluAl 1142
QY 3839 CATCGTTGGCAGCCCCCAGACCGATGATGATGACCAACGCCCGCATGACCCACTATCAAAG 3898
Db 1142 aLeuValArgGlnProProAsnLysTrpLeuSerAsnAlaArgMetHisTyrGlnAl 1162
QY 3899 CTTGCTTCTC---ACAGAGAGGTCACGTTTCGCTCCACCAGCCGCTCTCAACCTCGCAC 3955
Db 1162 aMetLeuLeuAspAlaGluArgValHisPheGlyProThrValSerLeuAsnProAlaTh 1182
QY 3956 TCTTCTGCTCAGAGAGACTGATGAACCA-----GTGACTCATGATTGCCATCAACTATT 4009
Db 1182 rLeuLeuProLeuProSerGlyLysProArgLeuSerProAsp----- 1197
QY 4010 GATTGAGGAGACTGGGGTCCGAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCT 4069
Db 1198 -LeuAlaGluThrMetAlaGlnThrAspLeuThrAspGlnProLeuProAspAlaAspLe 1217
QY 4070 AACCTGGTTCACTGACGGAAGCAGCTATGTTGGTGAAGGTAAAGAGATGCTGCGGCGGC 4129
Db 1217 uThrTrpTyrThrAspGlySerSerPheIleArgAsnGlyGluArgLysAlaGlyAlaAl 1237
QY 4130 GGTGTGAGACGGGACCCGACCATCTGGCGCAGCAGCTGCGCGGAAGGAACTTCAGACAC 4189
Db 1237 aValThrThrGluSerGluValIleTrpAlaAlaSerLeuProProGlyThrSerAlaGl 1257
QY 4190 AAAGCTGAGCTCATCGGCTCTACGCAAGCTTTGCGGCTGGCCGGAAGGAAATCCATAA 4249
Db 1257 nArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaLysGlyLysLeuTh 1277
QY 4250 CATTTATACGACAGGAGTATGCTTTCGCTGACCTGACACGTCACATATGGGCGCATATAA 4309
Db 1277 rValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyGluIleTyrAr 1297
QY 4310 ACAAGGGGGTTCCTTACCTCAGCAGGAGGGAATAAAGAAACAAAGAGAGAAATCTTAAG 4369
Db 1297 gArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAsnGluIleLeuAl 1317
QY 4370 CCTATTAGAGCCGTCATTTACCAAAAAGCTAGCTATTATACACTGTCTGCGACATCA 4429
Db 1317 aLeuLeuGluAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHisGl 1337

QY 4430 GAAAGCTAAGATCTCATATCCAGAGAAACACGATGGCTGACCGGGTTCCCAAGCAGGC 4489
 Db 1337 nlysglyAspSerProGlnAlaLysGlyAenArgLeuAlaAspThrAlaLysLysAl 1357
 QY 4490 AGCC-----CAGGGTGTTAACTTCTGCT-----ATAATGAATGCCAA 4531
 Db 1357 aAlaThrGluThrGlnSerSerLeuThrIleLeuProThrGluLeuIleGluGlyProly 1377
 QY 4532 AGCCCCAGAACCCAGACGACGATACACCCCTAGAGACTGGCAAGAGATAAAGATAGA 4591
 Db 1377 s-----ArgProProTrpGluTyArgAspSerAspLeuValGlnLysLeuG 1395
 QY 4592 CAGATTCTCGACATCCGGAAGCGGACCTCTATACCTCAGATGGGAAGAAATCTGCC 4651
 Db 1395 uAlaHisTyArgGluProLysArgGlyThr---TrpGluTyArgGlyLysThrIleMetPr 1414
 QY 4652 CCACAAAGAGGGTTAGATATATCCAAACAGATACATCGTCTAACCCACTAGGACTAA 4711
 Db 1414 oGluLysTyAlaLysGluLeuIleSerHisLeuHisLysLeuThrHisLeuSerAlaAr 1434
 QY 4712 ACACCTGCAGCAGTGTGGTCAGAACATCC-----CCTTATCATGTCT 4753
 Db 1434 glyMetLysThrLeuLeuGluArgGluGluThrGlyPheTyLeuProAsnArgAspLe 1454
 QY 4754 GAGGCTACACGAGTGGCTGACTCGGTGTTCAAACATTTGTGCTCCCTGCCAGCTGTTAA 4813
 Db 1454 uHisLeuArgGlnValThrGluSer-----CysArgAlaCysAlaGlnIleAs 1470
 QY 4814 TGCTAATCTTCCAGAAATGCTCCAGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCA 4873
 Db 1470 nAlaGlyLysIleLysPheGlyProAspValArgAlaArgGlyArgProGlyThrHi 1490
 QY 4874 CTGGGAAGTGGACTTCACCTGAGTAAAGCGCGCTAAATACGGAACCAATACCTATTGTT 4933
 Db 1490 strpGluValAspPheThrGluIleLysProGlyMetTyArgGlyTyLysTyLeuLeuVa 1510
 QY 4934 TTTTGTAGACACTTTTCAGGATGGGTAGAGCTTATCCTACTAAGAAAGAGACTTCAAC 4993
 Db 1510 lPheIleAspThrPheSerGlyTrpAlaGluAlaTyProAlaLysHisGluThrAlaLys 1530
 QY 4994 CTGGTGGCTAAATAATACTGGAAGAAATTTTCCAGATTTTGGATACCTTAAGTAAT 5053
 Db 1530 sValValAlaLysLysLeuLeuGluIlePheProArgTyArgIleProGlnValLe 1550
 QY 5054 AGGTCAGACAATGTGTCAGCTTTTGTGCCAGGTAAAGTCAGGAGCTGCCAAGATAT 5113
 Db 1550 uGlySerAspAsnGlyProAlaPheIleSerGlnValSerGlnSerValAlaThrLeuLe 1570
 QY 5114 GGGGATTGATTGAAACTGCTATGTCATACAGACCCCAAGCTCAGGACAGGTAGAGAG 5173
 Db 1570 uGlyIleAsnTrpLysLeuHisCysAlaTyArgProGlnSerSerGlyGlnValGluAr 1590
 QY 5174 GATGAATAGAACCATTAAGAGACCTTACTAATTTGACCGCGGAGACTGGGTAAATGA 5233
 Db 1590 gMetAsnArgSerIleLysGluThrLeuThrLysLeuThrLeuGluThrGlySerLysAs 1610
 QY 5234 TTGGATAGCTCTCTGCCCTTTGTGCTTTTGTAGGTAGAACACCCCTGGACAGTTGG 5293
 Db 1610 pTrpValLeuLeuLeuProLeuValLeuTyArgValArgAsnThrProGlyProHisG 1630
 QY 5294 GGTGACCCCTTAGAATTAATCTACGGGGGACCCCTTGGTAGAATTTGCTTCTGT 5353
 Db 1630 yLeuThrProPheGluIleLeuTyGlyAlaProProMetAlaHisPhePheAspTh 1650
 QY 5354 ACATAGTGTGAGTGTCTCTTCCAGCGCTTTGTCTTAGCTCAAGGCACTTGAGTG 5413
 Db 1650 rAspIleSerThrPheAlaThrSerProThrMetGlnAlaHisLeuArgAlaGlnLe 1670
 QY 5414 GGTGAGACAAACGAGGTGGAGCAACTCCGGGAGGCTACTCAGGAGGAGGAGCTTGA 5473
 Db 1670 uValGlnGluGluIleGlnArgProLeuAlaAlaTyArgGluLysLeuGluThrPr 1690
 QY 5474 G---ATCCACATCGTTTCCAAAGTGGGAGATTCACTTACGTTAGACGCGCACCGTGCAGG 5530

Db 1690 oValValProHisProPheLysProGlyAspSerValTrpValArgArgHisGlnThrLy 1710
 QY 5531 ARACTCTGAGACTCGGTGGAGGGCCCTTATCTGCTACTTTTGCACACACCAAGCGTGT 5590
 Db 1710 sAnLeuGluProArgTrpLysGlyProHisIleValLeuLeuThrThrProThrAlaLe 1730
 QY 5591 GAAAGTCGAAGAATCTCCACCTGGATCCATGATCCCATGTTAAACCGCGCCACCT-- 5648
 Db 1730 uLysValAspGlyValAlaAlaTrpIleHisAlaSerHisValLysAlaAlaGlyProTh 1750
 QY 5649 -----CCCGATTCCGGGTGGAAAGCCGA 5671
 Db 1750 rThrAsnGlnAspLeuSerAspSerProSerSerAspAspProSerArgTrpLysValG 1770
 QY 5672 AAGACTGAAATCCCTTAAAGCTTCGCTCCATCGC 5708
 Db 1770 nArgThrGlnAsnProLeuLysIleArgLeuSerArg 1782
 RESULT 7
 AAR05898
 ID AAR05898 standard; protein; 1784 AA.
 XX
 AC AAR05898;
 XX
 DT 24-OCT-2003 (revised)
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 29-NOV-1990 (first entry)
 XX
 DE Gene product of first ORF of virus capable of inducing viraemia in feline species encoded by clone 61E.
 XX
 KW Viraemia; leukaemia; FeLV-A; clone 61C; clone 61E; ds.
 XX
 OS Feline leukemia virus; strain A.
 XX
 PN EP377842-A.
 XX
 PD 18-JUL-1990.
 XX
 PF 12-DEC-1989; 89EP-00122964.
 XX
 PR 13-DEC-1988; 88US-00284139.
 XX
 PA (HARD) UNIV HARVARD
 PA (COLS) COLORADO STATE UNIV RES.
 PA (HARD) UNIV HARVARD.
 XX
 PI Hoover EA, Mullins JI;
 DR WPI: 1990-218326/29.
 DR N-PSDB; AAQ05252.
 XX
 PT DNA encoding leukaemia virus-A sub: type - producing AIDS type disease in cats, used to test drugs and vaccines.
 XX
 PS Claim 6; Fig 1.1-1.7; 23pp; English.
 XX
 CC Variant of viral genome induces immunodeficiency in cats similar to AIDS. May be used in research, especially in testing drugs and vaccines against viraemia and feline leukaemia viruses. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 1784 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 1784
 Score: 5229.00 Matches: 1028
 Percent Similarity: 70.3% Conservative: 235
 Best Local Similarity: 57.2% Mismatches: 410
 Query Match: 35.7% Indels: 123

DB:	2	Gaps:	29
US-10-723-552-3 (1-8132) x AAR05898 (1-1784)			
QY	512	TCACGTGTGCTGCATCTGTGGTTCTGTTTGTGTGTCCTTTGTCTGTGTGTCCTTGT	571
Db	47	SerArgAlaAlaArgLeuValIlePheCysLeuValAlaSerPheLeuVal---ProCys	65
QY	572	CTACAGTTTT-----AATATGGGACAGCGGTACGACCCCTCTTACGTTT	616
Db	66	Leu-ThrPheLeuIleAlaGluThrValMetGlyGlnThrIleThrThrProLeuSerLe	85
QY	617	GACTCTCGACCATTTGAGCTGAAGTTAAATCCAGGGCTCATAAATTTGCAGTTCAGGTTAA	676
Db	85	uThrLeuAspHisTrpSerGluValArgAlaArgAlaHisAsnGlnGlyValGluValAr	105
QY	677	GAAGGGACCTTGGCAGACTTCTGTGTCCTCTGAATGCCGACATTCATGTTGGATGCC	736
Db	105	gLyLeuLysTrpIleThrLeuCysGluAlaGluTrpValMetMetAsnValGlyTrpPr	125
QY	737	ATCAGAGGGACCTTAAATCTGAGATATCTCGCTGTTAAAGCAGTATTTTTCAGAC	796
Db	125	oArgGluGlyThrPheSerLeuAspAsnIleSerGlnValGluLysIlePheAlaPr	145
QY	797	TGGACCCGGCTCTCATCCGATCAGGAGCCCTATATCTTACGTGGCAAGATTTGGCAGA	856
Db	145	oGlyProTyrGlyHisProAspGlnValProTyrIleThrThrTrpArgSerLeuAlaTh	165
QY	857	GGATCTCCGCCATGGGTTAAACCATGGCTGAATAAGCAAGAAAGCAGGTCCCCGAAT	916
Db	165	rAspProProSerTrpValArgProPheLeuProProLysThrProThrPro-----	183
QY	917	TCTGGCTCTGGACAGAAAAACAACACTCGGCTGAAAAAGTCAAGCCCTCTCCT---	971
Db	184	-----LeuProGlnProLeuSerProGlnProSerAl	194
QY	972	-----CATATCTACCCCGAGATT-----GAGGAGCCACCGCGC	1003
Db	194	aProLeuThrSerSerLeuTyTrpValLeuProLysSerAspProProLysProProVa	214
QY	1004	TTGGCCCGGAACCCCAATCTGTTCCT-----CCACCCCTCTTA	1039
Db	214	lLeuProProAspProSerProSerProLeuIleAspLeuLeuThrGluGluProProTy	234
QY	1040	TCTGGCACAGGTCCTCGGAGGACCTTTGCCCTCTCGGACTCCG-----GCGGT	1093
Db	234	r-----ProGlyGlyHisGlyPro-----ProProSerGlyProArgThrProTh	249
QY	1094	GGAGGACCTCTGCAGGAGACTCGGAGCCGGAGGGGGCCCAACCCGGAGCGGACAGACA	1153
Db	249	rAlaSerProIleAlaSerArgLeuArgGluArgGluAsnPro-----AlaGluGl	267
QY	1154	GATCGCGACATTACCGCTCGCAGCTAGGGCCCTCCACACCGGGGGGCCAATTGACGCC	1213
Db	267	uSerGlnAlaLeuProLeuArg-----GluGlyProAsnAsnArgPr	281
QY	1214	CCTCCAGTATTGGCCCTTTCTTCTGCGAGACTCTATAATTGGAAAACTAACATCCCC	1273
Db	281	o-----GlnTyTrpProPheSerAlaSerAspLeuTyrAsnTrpLysSerHisAsnProPr	300
QY	1274	TTTTCTCGAGGATCCCCAAACCCCTCAGGGGTGGTGGAGTCCCTTATGTTCTCTCACCA	1333
Db	300	oPheSerGlnAspProValAlaLeuThrAsnLeuIleGluSerIleLeuValThrHisGl	320
QY	1334	GCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTTCACAAACCGAGGAGCGAGA	1393
Db	320	nProThrTrpAspAspCysGlnGlnLeuLeuGlnAlaLeuLeuThrGlyGluGluArgGl	340
QY	1394	GAGAAITCTATTAGAGCTAGAAAATGTTCTCGGGCCGACCGGCGCAGCCCGCGGTT	1453
Db	340	nArgValLeuLeuGluAlaArgLysGlnValProGlyGluAspGlyArgProThrGlnLe	360
QY	1454	GCAAAATGAGATTGACATGGGATTTCCCTTAACCTCGCCCGCGGTGGGACTACAAACGCGC	1513

Db	360	uProAsnValileAspGluAlaPheProLeuThrArgProAsnTrpAspPheArgThrPr	380
Qy	1514	TGAAGGTAGGAGAGCTTGAATAATCTATCCAGGCTCTGGTGGCGGTCTCGGGCGC	1573
Db	380	oAlaGlyArgGluHisLeuArgLeuThrArgGlnLeuLeuLeuAlaGlyLeuArgGlyAl	400
Qy	1574	CTCAAGACGGCCCACTAATTTGGCTAAGTAAGAGAGTGTATCAGGAGCCGCAATGAACC	1633
Db	400	aAlaArgArgProThrAsnLeuAlaGlnValLysGlnValValGlnGlyLysGluGluTh	420
Qy	1634	CCCTCTGTTTTTCTTGAGAGGCTCTTGGAAAGCTTTCAGGCGGTACACCCCTTTTGTATCC	1693
Db	420	rProAlaSerPheLeuGluArgLeuLysGluAlaTyrArgMetTyrThrProTyrAspPr	440
Qy	1694	CACCTCAGAGCGCCCAAAAGCCCTCAGTGGCTTTGGCCCTTTATAGGACAGTCAGCTTGGGA	1753
Db	440	oGluAspProGlyGlnAlaIleSerValIleLeuSerPheIleTyrGlnSerSerProAs	460
Qy	1754	TATTAGAAAGAGCTTCAGAGACTGGAAAGGTTTACAGAGGCTCAGTTACTGATCTAGT	1813
Db	460	pIleArgAsnLysLeuGlnArgLeuGluGlyLeuGlnGlyPheThrLeuSerAspLeuLe	480
Qy	1814	GAAGGAGCGCAGAAAGTATATTACAAAAGGAGACAGAAAGAGAAAGGAAACAAGAAA	1873
Db	480	uLysGluAlaGluLysIleTyrAsnLysArgGluThrProGluGluArgGluGluArg--	499
Qy	1874	AGAGAGAGAAGAGAGCAAGGAGGAGAAACACGTATAATAACGCAAGAGAAAGATTTCAC	1933
Db	500	-----LeuTrpGlnArgGlnGluGluArgAspLysIleArgHisLysGluMetTh	516
Qy	1934	TAAGATCTTGCTGCAGTGGTTCAAGGGAAAAACATACGCAAGAGACAGAGA-----	1985
Db	516	rLysValLeuAlaThrValValAlaGlnAsnArgAspLysAspArgGluGluSerLysLe	536
Qy	1986	----GATTTAGGAAATTAGTTCAGGCCCTAGACAGCTCAGGGAACCTGGCAATAGGAC	2041
Db	536	uGlyAspGlnArgLysIle-----	542
Qy	2042	CCCACTCGACAGGACCAATGTGCATATTCTAAGAAGAGGACACTGGCAAGGACTG	2101
Db	543	-ProLeuGlyLysAspGlnCysAlaTyrCysLysGluLysGlyHisIleTrpValArgAspC	562
Qy	2102	CCCCAAGAGGGAACAAAGACCA-----AGGATCCTAGCTCTAGAAGAAAGATAA	2152
Db	562	sProLysArgProArgLysLysProAlaAsnSerThrLeuLeuAsnLeuGluAsp***G	582
Qy	2153	AGATTAGGAGAGCGGGTTCGGACCCCTCCCGAGCCCAAGGTAACCTTTGAAGTGGGA	2212
Db	582	user-----GlnGlyGlnAspProProGluProArgIleThrLeuLysIleG	599
Qy	2213	GGGCAACCACTGAGTCTCTGGTTGATACCGGAGCGAAACATTTCAGTGTCTACTACAGCC	2272
Db	599	yGlyGlnProValThrPheLeuValAspThrGlyAlaGlnHiserValLeuThrArgPr	619
Qy	2273	ATTAGGAAACATAAAGATAAAAAATCCTGGGTGATGGGTGCCACAGGCAACAAACAGTA	2332
Db	619	oAspGlyProLeuSerAspArgThrAlaLeuValGlnGlyAlaThrGlySerLysAsnTy	639
Qy	2333	TCCATGAGCTACCCGAAGAACAGTTGACTTGGGAGTGGGACCGGATAACCCACTCGTTTCT	2392
Db	639	rArgTrpThrThrAspArgValGlnLeuAlaThrGlyLysValThrHiserPheLe	659
Qy	2393	GGTCATACCTGAGTCCCGACGACCCCTCTTAGGTAGAGACTTATTCAGCAGATGGGAGC	2452
Db	659	uTyValProGluCysProTyrProLeuLeuGlyArgAspLeuLeuThrLysLeuLysAl	679
Qy	2453	ACAAATTTCTTTTGA--CAAGGGAAACAGAAAGTGTCTGCAAAATAACAAACCTATCAC	2509
Db	679	aglnIleHisPheThrGlyGluGlyAlaAsnValValGlyProArgGlyLeuProLeuG	699
Qy	2510	TGTTGTCACCTCCAAATTAGATACGAAATATCGACTATATCTCTCCCTAGTAAAGCCTCGA	2569

Db 699 nValLeuThrLeuGlnLeuGluGluTyrArgLeuPheGluProGluSerThrGlnIly 719
QY 2570 TCAAAATATACAAATCTGGTGTGAACAGTTTCCCAAGCCTCGGCGAGAAACCCAGCGGAT 2629
Db 719 eGlnGluMetAspIleTrpLeuLysAsnProGlnAlaTrpAlaGluThrGlyGlyMe 739
QY 2630 GGGTTTGGCAAGCAAGTCCCCCAACAGTATTTCACAGGCGCAGTGCACACAGT 2689
Db 739 tGlyMetAlaHisCysGlnAlaProValLeuIleGlnLeuLysAlaThrAlaThrProIle 759
QY 2690 GTCAAGTCAGACACTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCGCATGTCCA 2749
Db 759 eSerIleArgGlnTyrProMetProHisGluAlaTyrGlnGlyIleuProHisIleAr 779
QY 2750 AGATTAAATCAACAGGCGCATCTAGTTCCTGTCCTCAATCTCCCTGGAATACTCCCTGCT 2809
Db 779 gArgMetLeuAspGlnGlyIleLeuLysProCysGlnSerProTrpAsnThrProLeuLe 799
QY 2810 ACCGGTTAGAAAGCCTGGAGCTAATGACTATCCACAGTACAGGACTTGAGAGAGTCAA 2869
Db 799 uProValLysLysProGlyThrLysAspTyrArgProValGlnAspLeuArgGluValAs 819
QY 2870 TAAACGGGTGCAGATATACACCAACAGTCCGGAACCTTATAACCTCTGTGTCTCT 2929
Db 819 nLysArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerThrLe 839
QY 2930 CCCACCCCAACGAGCTGGTATACAGTATTGGACTTAAAGGATGCCTTCTTCGCTGAG 2989
Db 839 uProProSerHisProTrpTyrThrValLeuAspLeuLysAspAlaPheCysLeuAr 859
QY 2990 ATTAACCCCACTAGCAACCACTTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAG 3049
Db 859 gLeuHisSerGluSerGlnLeuLeuPheAlaPheGluTrpArgAspProGluIleGlyLe 879
QY 3050 AACCGGGAGCTCACCTGACCGACTGCGCCCAAGGGTTCAGAACTCCCGACCATCTT 3109
Db 879 uSerGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPh 899
QY 3110 TGACGAAGCCCTACACAGAGACCTGGCCCAACTTCAGAGTCCACACCTCAGGTGACCCT 3169
Db 899 eAspGluAlaLeuHisSerAspLeuAlaAspPheArgValArgTyrProAlaLeuValle 919
QY 3170 CTTCCAGTACGTGGATGACTGCTTCTGGCGGAGCCCAACACAGGACTGCTTGAAGG 3229
Db 919 uLeuGlnTyrValAspAspLeuLeuAlaAlaAlaThrArgThrGluCysLeuGluGl 939
QY 3230 CACGAAGGCACTACTGCTGAATGTCTGACCTAGCTACAGGCTCTGCTAAGAGGC 3289
Db 939 yThrLysAlaLeuLeuGluThrLeuGlyAsnLysGlyTyrArgAlaSerAlaLysLysAl 959
QY 3290 CCAGATTTTCAGGAGAGAGTAACTACTTGGGGTACAGTTTGGCGGACGGCGCAGCATG 3349
Db 959 aGlnIleCysLeuGlnGluValThrTyrLeuGlyTyrSerLeuLysAspGlyGlnArgTr 979
QY 3350 GCTGACGAGGCGCAGGAAGAAACTGTATGTCAGATACCGGCGCCCAACACAGCCAAACA 3409
Db 979 pLeuThrLysAlaArgLysGluAlaIleLeuSerIleProValProLysAsnProArgI 999
QY 3410 RATGAGAGATTTTGGGACAGCTGATTTTGCAGACTGCGATCCCGGGGTTCGAC 3469
Db 999 nValArgGluPheLeuGlyThrAlaGlyTyrCysArgLeuTrpIleProGlyPheAlaGl 1019
QY 3470 CTTAGCAGCCCACTCTACCCCTTAACCAAGAAAAGGGGAATTCCTCGGGCTCCTGA 3529
Db 1019 uLeuAlaAlaProLeuTyrProLeuThrArgProGlyThrLeuPheGlnTrpGlyThrGl 1039
QY 3530 GCACAGAGGCAATTTGATGCTATCAAAAGGCCCTGCTGAGCGCACCTGCTCGGCCCT 3589
Db 1039 uGlnGlnLeuAlaPheGluAsnIleArgLysAlaLeuLeuSerSerProAlaLeuGlyLe 1059
QY 3590 CCTCAGCTAACTAAACCTTTACCTTTATGTGATGAGCTAAGGAGTAGCCCGGG 3649
Db 1059 uProAspIleThrLysProPheGluLeuPheIleAspGluAsnSerGlyPheAlaLysGl 1079

QY 3650 AGTTTTTAACCCAAACCTAGGACCATGGAGAACCTGTGCGCTACTCTGTCAAAGAGCT 3709
Db 1079 yValLeuValGlnLysLeuGlyProTrpLysArgProValAlaTyrLeuSerLysLysLe 1099
QY 3710 CGATCTGTAGCAGTGGTGGCCCATATGCTTGAAGGCTATCGAGCTGTGTGCCATACT 3769
Db 1099 uAspThrValAlaSerGlyTrpProCysLeuArgMetValAlaAlaIleLe 1119
QY 3770 GGTCAAGGACGCTGACAAATTCAGCTTTGGACAGATATAACTGTAAATAGCCCCCATGC 3829
Db 1119 uValLysAspAlaGlyLysLeuThrLeuGlyGlnProLeuThrIleLeuThrSerHisPr 1139
QY 3830 ATTGGAGAACTCGTTTCGGCAGCCCCCAGACCATGGATGACCAACGCGCCGATGACCCA 3889
Db 1139 oValGluAlaLeuValArgGlnProProAsnLysTrpLeuSerAsnAlaArgMetThrHi 1159
QY 3890 CTATCAAGCCTGCTTCTC---ACAGAGAGGTTCAGTTCGCTCCACACGCCCTCTCAA 3946
Db 1159 sTyrGlnAlaMetLeuLeuAspAlaGluArgValHisPheGlyProThrValSerLeuAs 1179
QY 3947 CCTGCGCACTCTTCTGCTGGAAGAGACTGATGAACCA-----GTGACTCATGATGCCA 4000
Db 1179 nProAlaThrLeuLeuProLeuProSerGlyLysProProArgLeuSerProAsp---- 1197
QY 4001 TCAACTATTGATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGG 4060
Db 1198 -----LeuAlaGluThrMetAlaGlnThrAspLeuThrAspGlnProLeuProAs 1214
QY 4061 AGAAGTGTAACTCGTGTCTACGCGGAAGCAGCTATGTGGTGAAGAGTGAAGAGGATGCC 4120
Db 1214 pAlaAspLeuThrTrpTyrThrAspGlySerSerPheIleArgAsnGlyGluArgLysAl 1234
QY 4121 TCGGGCGGGTGGCGGACCGCACGATCTGGGCGCAGCAGCTCCCGGAGGAGAAC 4180
Db 1234 aGlyAlaAlaValThrThrGluSerGluValIleTrpAlaAlaSerLeuProProGlyTh 1254
QY 4181 TTCAGCAAAAAGGCTGAGCTCATGGCCCTCACGCAAGCTTTTGGCGCTCGCGCAAGCGAA 4240
Db 1254 rSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaLysGlyLy 1274
QY 4241 ATCCATAAACAATTTATAGCAGCAGAGTATGCTTTGGAGTGCACACGTACATCGGGC 4300
Db 1274 sLysLeuThrValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyGl 1294
QY 4301 CATCTATAAACAAGGGGTGCTTACTCTCAGCAGGAGGGAATAAAGAACAAGAGGA 4360
Db 1294 uIleTyrArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAsnGl 1314
QY 4361 AATTTAAGCCTATTAGAACCGCTACATTTACCAAAAAGGCTAGCTATTATACACTGTCC 4420
Db 1314 uIleLeuAlaLeuLeuGluAlaLeuPheLeuProLysArgLeuSerIleIleHisCysPr 1334
QY 4421 TGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATGGCTGACCGGGTGC 4480
Db 1334 oGlyHisGlnLysGlyAspSerProGlnAlaLysGlyAsnArgLeuAlaAspThrAl 1354
QY 4481 CAAGCAGCAGCC-----CAGGGTCTTAACCTTCTGCTCCT-----ATATAGA 4522
Db 1354 aLysLysAlaAlaThrGluThrGlnSerSerLeuThrIleLeuProThrGluLeuIleGl 1374
QY 4523 AATGCCAAAAGCCCGAGAACCCAGACGACGATACACCTTAGAAGACTGGCAAGAGATAA 4582
Db 1374 uGlyProLys-----ArgProProTrpGluTyrAspAspSerAspLeuAspLeuValGl 1392
QY 4593 AAAGATAGACCAATTTCTCTGAGACTCCGGAAGGGACCTGCTATACCTCAGATGGGAAGA 4642
Db 1392 nLysLeuGluAlaHisTyrGluProLysArgGlyThr---TrpGluTyrArgGlyLysTh 1411
QY 4643 AATCTGCCCCCAAGAAAGGGTTAGATATGTCCACAGATACATCGTCTAACCCACCT 4702
Db 1411 rIleMetProGluLysTyrAlaLysGluLeuIleSerHisLeuHisLysLeuThrHisLe 1431

Qy	4703	AGGAACCTAAACACCTCGACGACGAGTTGGTCAGACAATCC-----CCTTA	4744
Db	1431	userAlaArgLysMetLysThrLeuLeuGluuArgGluGluThrGlyPheTyrLeuProAs	1451
Qy	4745	TCATGTTCTGAGGCTACCGAGAGTGGCTGACTGGTGGTCANACATTTGTGTGCCCTGCCA	4804
Db	1451	nArgAspLeuHisLeuArgGlnValThrGluSer-----CysArgAlaCysAl	1467
Qy	4805	GCTGGTTAATGCTTAATCTCTCCAGAAATGCCTCCAGGAGAGAGACTAAGGGGAAGACACCC	4864
Db	1467	gAlnIleAsnAlaGlyLysIleLysPheGlyLeuAspValArgAlaArgGlyArgArgPr	1487
Qy	4865	AGGCGCTACTGGGAAGTGGACTTCACTAGGTAAACCGCGCTAAATACCGAAACAATA	4924
Db	1487	oGlyThrHisTrpGluValAspPheThrGluIleLysProGlyMetTyrGlyTyrLysTy	1507
Qy	4925	CCTATTGGTTTGTAGACACCTTTTCAGAGTGGGTAGAGCTTATCTACTAAGAAAGA	4984
Db	1507	rLeuLeuValPheIleAspThrPheSerGlyTrpAlaGluAlaTyrProAlaLysHisGln	1527
Qy	4985	GACTTCAACCGTGTGGCTAAAAAATACTGGAAGAAATTTTCCAGATTTCCGAATACC	5044
Db	1527	uThrAlaLysValAlaLysLysLeuGluGluIlePheProArgTyrGlyIlePr	1547
Qy	5045	TAAGTAAATAGGGTCAGACAAATGGTCCAGCTTTTGTTCGCCAGTAAGTCAGGACTGGC	5104
Db	1547	oGlnValLeuGlySerAspAsnGlyProAlaPheIleSerGlnValSerGlnSerValAl	1567
Qy	5105	CAAGATATTGGGGATTGATGGAAACTGCATTTGTGCATACAGACCCCAAGACTCAGGACA	5164
Db	1567	aThrLeuLeuGlyIleAsnTrpLysLeuHisCysAlaTyrArgProGlnSerSerGlyGln	1587
Qy	5165	GGTAGAGAGATGAATAGAACCATTAAAGAGACCTTACTAAATTGACCGCGAGACTGG	5224
Db	1587	nValGluArgMetAsnArgSerIleLysGluThrLeuThrLysLeuThrLeuGluThrGln	1607
Qy	5225	CGTTAAATGATTGGATAGCTCTCTCGCCCTTTGTGCTTTTATGAGGTAGGAACACCCCTGG	5284
Db	1607	ySerLysAspTrpValLeuLeuLeuProLeuValLeuTyrArgValArgAsnThrProGln	1627
Qy	5285	ACAGTTTGGCTGACCCCTATGAATTACTCTACGGGGGACCCCCCCCATTTGTTAGAAAT	5344
Db	1627	yProHisGlyLeuThrProPheGluIleLeuTyrGlyAlaProProProMetAlaHisPh	1647
Qy	5345	TGCTTCTGTACATAGTCTGACGTGCTGCTTCCAGCCCTTTGTTCTCTAGGCTCAAGGC	5404
Db	1647	ePheAspThrAspIleSerThrPheAlaThrSerProThrMetGlnAlaHisLeuArgAl	1667
Qy	5405	ACTTGAGTGGGTGAGACAACAGCGCTGGAGGCAACTCCGGGAGCCCTACTCAGAGGAGG	5464
Db	1667	aLeuGlnLeuValGlnGluGluIleGlnArgProLeuAlaAlaTyrArgGluLysLe	1687
Qy	5465	AGACTTGCAG---ATCCACATCGTTTCCAGTGGGAGATTCACTTACGTTAGACGCCA	5521
Db	1687	uGluThrProValValProHisProPheLysProGlyAspSerValTrpValArgArgHis	1707
Qy	5522	CCGTGCAGGAAACCTCGAGACTGGTGGGAAGGGCCCTTACTCGTACTTTTGCACACACC	5581
Db	1707	sGlnThrLysAsnLeuGluProArgTrpLysGlyProHisIleValLeuLeuThrThrPr	1727
Qy	5582	AACGGCTGTGAAGTCCGAAGAAATCTCCACTGGATCCATGCATCCACAGCTTAAACCGGC	5641
Db	1727	oThrAlaLeuLysValAspGlyValAlaAlaIlePheHisAlaSerHisValLysAlaAl	1747
Qy	5642	GCCACCT-----CCCGATTTCGGGGTG	5662
Db	1747	aGlyProThrThrAsnGlnAspLeuSerAspSerProSerSerAspAspProSerArgTr	1767
Qy	5663	GAAGCCGAAAAGACTGAAAATCCCTTAACTTCGCTCGCTCATCGC	5708
Db	1767	pLysValGlnArgThrGlnAsnProLeuLysIleArgLeuSerArg	1782

AAB73284	
ID	AAB73284 standard; protein; 1069 AA.
XX	
AC	AAB73284;
XX	
DT	23-MAY-2001 (first entry)
XX	
DE	Defective retroviral genome protein #3 isolated from PK-15 cell line.
XX	
KW	Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line.
XX	
OS	Unidentified.
XX	
PN	US6190861-B1.
XX	
PD	20-FEB-2001.
XX	
PF	13-DEC-1996; 96US-00766528.
XX	
PR	14-DEC-1995; 95US-00572645.
XX	
PA	(GHO) GEN HOSPITAL CORP.
XX	
PI	Fishman JA;
XX	
DR	WPI; 2001-256211/26.
XX	
DR	N-PSDB; AAF77726.
XX	
AS	Assessing risk of endogenous retroviruses in clinical practice and in
PT	xenotransplantation, comprises using probe sequences derived from swine
PT	or miniature swine retroviral genome.
XX	
PS	Disclosure; Fig 2; 127pp; English.
XX	
CC	The present invention relates to a method for screening a cell or tissue
CC	for the presence or expression of a retrovirus (RV), comprising
CC	contacting a target nucleic acid from the cell or tissue with a second
CC	nucleic acid from the present invention (e.g. AAF77726 or a fragment
CC	thereof). The method is useful for RV detection and to assess graft
CC	transplantation risk. Screening of animals allows the elimination of
CC	donors with active replication of known viruses. Inactive proviruses can
CC	be detected and inactivated, allowing identification and elimination of
CC	potential human pathogens derived from swine in a manner not possible in
CC	the outbred human organ donor population and is important to the
CC	development of human xenotransplantation
XX	
SQ	Sequence 1069 AA;
Alignment Scores:	
Pred. No.:	0 Length: 1069
Score:	4889, 00 Matches: 930
Percent Similarity:	89.8% Conservative: 46
Best Local Similarity:	85.6% Mismatches: 62
Query Match:	33.4% Indels: 49
DB:	4 Gaps: 8
US-10-723-552-3 (1-8132) x AAB73284 (1-1069)	
QY	4311 CAAGGGGGTGTCTTACCTCAGCAGGAGGGAATAAGAACAAAGAGGAAATTC TAAGC 4370
	:::
Db	1 LysArgGlyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSer 20
QY	4371 CTATTAGACCGGTACATTTTACCAAAAGGCTAGCTATTATATACACTGTCTGTGACATCAG 4430
	:::
Db	21 LeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGln 40
QY	4431 AAAGCTAAAGATCTCATATCCAGAGGAACACAGATGGCTCAGCCGGTTCACAAGCAGGCA 4490
	:::
Db	41 LysAlaLysAspLeuIleSerArgLysGlnMetAlaAspArgValAlaLysGlnAla 60
QY	4491 GCCCAGGGTGTAACTTCTGCCTTATTAATAGAAATGCCAAAGCCCAAGCCAGACGA 4550
	:::
Db	61 AlaGlnAlaValAsnLeuLeuProIleIleGluThrProLysAlaProGluProAlaArg 80

QY 4551 CAGTACACCTAGAGACTGGCAGAGATATAAAGATAGACCACTGTTCTCTGAGACTCCG 4610
 Db 81 GlnTyrThrLeuGluAseTrpGlnGluIleLysLysIleAsePheSerGluThrPro 100
 QY 4611 GAAGGACCTGCTATACCTCAGATGGAGGAATCTGCGCCACAAAGAGGGTTAGAA 4670
 Db 101 GluGlyThrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGlu 120
 QY 4671 TATGTCACACAGATACATCTCTAACCCACCTTAGGAACATAACACCTGCAGCAGTGGTC 4730
 Db 121 TyrValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnLeuVal 140
 QY 4731 AGAACATCCCTTATCATGTTCTGAGCTACACGAGTGGCTGACTCGGTGGTCAACAT 4790
 Db 141 ArgThrSerProTyrHisValLeuArgLeuProGlyValAlaAseSerValLysHis 160
 QY 4791 TGTGTCCTCCCTGCGCTGTTTAACTGCTTAATCTCCAGATGCTCCAGGAAGACTA 4850
 Db 161 CysValProCysGlnLeuValAsnAlaAsnProSerArgIleProGlyLysArgLeu 180
 QY 4851 AGGGAGGCCACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAA 4910
 Db 181 ArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLys 200
 QY 4911 TAGCGAAACAAATACCTATTGGTTTGTGTAGACACCTTTTTCAGGATGGGTAGGCTTAT 4970
 Db 201 TyrGlyAsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyr 220
 QY 4971 CCTACTAAGAAAGACTTCAACCGTGGTGGCTTAAATAAATACTGGAAGAAATTTTCCA 5030
 Db 221 ProThrLysLeuGluThrSerThrValValAlaLysLysIleLeuGluIlePhePro 240
 QY 5031 AGATTTGGAATACCTAAGGTATAGGTGTCAGACAAATGGTCCAGCTTTTGTGCCAGGTA 5090
 Db 241 ArgPheGlyLeuProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnVal 260
 QY 5091 AGTCAGGAGTGGCCAGATATTGGGATTTGATGGAAACTGCAATTTGTCATACAGACCC 5150
 Db 261 SerGlnGlyLeuAlaLysIleLeuGlyIleAse**LysLeuHisCysAlaTyrArgPro 280
 QY 5151 CAAGCTCAGGACAGTATAGAGGATGAATAGAACCATTAAGAGACCTTACTAATATG 5210
 Db 281 GlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeu 300
 QY 5211 ACCGCGAGACTGGCTTAATGATGATAGTCTCTCTGCTGCTGCTGCTTTTGGGTT 5270
 Db 301 ThrThrGluThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgVal 320
 QY 5271 AGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAATTTACTTACGGGGGACCCCC 5330
 Db 321 ArgAsnThrProGlyGlnPheGlyLeuThrProTyrLysLeuLeuTyrGlyProPro 340
 QY 5331 CCAATGGTGAATTCCTCTGTATAGTGTGAGTGTGCTGCTTCCAGCTTTGTTTC 5390
 Db 341 ProLeuAlaGluIleAlaPheAlaHisSerAlaAseValLeuLeuSerGlnProLeuPhe 360
 QY 5391 TCTAGGCTCAAGCACTCAGTGGGTGAGACACAGCGTGGAGCAACTCCGGGAGGCC 5450
 Db 361 SerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAla 380
 QY 5451 TACTCAGGAGGAGGAGACTTGAGATCCACATCGTTTCCAAAGTGGGAGATTCAGTCTAC 5510
 Db 381 TyrSer--GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyr 399
 QY 5511 GTTAGCGCACCGTGCAGAAACCTCGAGACTCGGTGGAAGGCCCTTATCTGACTT 5570
 Db 400 ValArgArgHisArgAlaGlyAsnLeuGluThrArg**LysGlyProTyrLeuValLeu 419
 QY 5571 TTGACCACACCAAGCGCTGGAAGTCAAGGAATCTCCACCTGGATCCATGCATCCAC 5630
 Db 420 LeuThrThrProThrAlaValLysValGluGlyIle-Pro----- 432

QY 5631 GTTAAACCGCGCCACCTCCCGATTCCGGGTGGAAGCCGAAAGACTGAAATCCCTT 5690
 Db 433 -----Le 433
 QY 5691 AAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTCAATAACCTCTCAGACTTAATGGTAT 5750
 Db 433 userPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGlnValAsnGlyLys 453
 QY 5751 GCCATAGGAGACAGCTGAACCTCCATAAACCTTATCTCTCACCTGGTAAATTAATCTGA 5810
 Db 453 sArgLeuValAseSerProAsnSerHisLysProLeuSerLeuThrTrpLeuLeuThrAs 473
 QY 5811 CTCGGCACAGGTATTAAATATCAACACACTCAAGGGGAGGCTCTTTTAGAACCTGGTG 5870
 Db 473 pSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeuGlyThrTrpTr 493
 QY 5871 GCCTGATCTATAGCTTGGCTCAGATCAGTTATTCCTAGTCTG-----ACCTCACC 5921
 Db 493 pProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAspGlnAlaThrPr 513
 QY 5922 CCAGATATCTCCATGCTCACGATTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5981
 Db 513 oProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProProAsnAsnGluGl 533
 QY 5982 ACATTCGCGAAATCCAGAGATTTCTTTTAAACAATGGAACCTGTAACTCTAATGA 6041
 Db 533 uTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysIleThrSerAsnAs 553
 QY 6042 TGGATATTGGAATGCCAACCTCTCAGCAGGATAGGTAAAGTTTCTTATGTCAACAC 6101
 Db 553 pGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSerPheValAsnAs 573
 QY 6102 CTATACCACTCTGGACAATTTAATTAC-----CTGACCTGG----- 6138
 Db 573 nProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAspTrpGlnGlnAr 593
 QY 6139 -----ATTAGAACTGGAAGCCCAAGTCTCTCTCTCAGACCTAGATTACCT 6185
 Db 593 gValGlnLysAseValAlaArgAsnLysGlnIleSerCysHisSerLeuAspLeuAspTyrLe 613
 QY 6186 AAAATAAGTTTCACTGAGAAAGGAAACAGAAATATCTTAAATGGGTAAATGGTAT 6245
 Db 613 uLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrpValAsnGlyL 633
 QY 6246 GTCTGGGAAATGGTATATTATGGAGGCTCGGTAAACACCAAGGCTCCATCTCACTAT 6305
 Db 633 eSerTrpGlyIleValTyrTyrGlyLysSerGlyArgLysLysLysLysLysLysLysL 653
 QY 6306 TCGCTCTCAAATA---AACGAGTGGAGCTCCAATGGCTATAGGACCAAAATACGGTCTT 6362
 Db 653 eArgLeuArgIleGluThrGlnMetGluProValAlaIleGlyProAsnLysGlyLe 673
 QY 6363 GACGGGTCAAGACCCCAACCCAA-----GGACGAGGACCATCTCT---AA 6407
 Db 673 uAlaGluGlnGlyProProIleGlnGluArgProSerProAsnProSerAspTyrAs 693
 QY 6408 CATAACTCTCGATCAGACCCCACTGAGTCAACAGCAGCTAAATGGGCGCAAACT 6467
 Db 693 nThrThrSerGlySerValProThrGluProAsnIleThrIleLysThrGlyAlaLysLe 713
 QY 6468 TTTTAGCTCATPCAGGAGGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAGGCTACCTC 6527
 Db 713 uPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThrSe 733
 QY 6528 TTCTTTGGCTATCTTAGCTTCGGGCCACCTTACTATGAAAGGATCGGTAGAGAGG 6587
 Db 733 tSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMetAlaArgGlyGl 753
 QY 6588 GAAATTCATGTGACAAAAGACATAGAGACCAATGCATCGGATCCCAAAATAAGCT 6647
 Db 753 yLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLysLe 773
 QY 6648 TACCCTTACTGAGGTTTCTGGAAAAGGACCTGCATAGGAAAGGTTTCCCCCATCCACCA 6707

Db 773 uThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValProSerHisGI 793
QY 6708 ACACCTTTGTAACACACACTGAAGCTTTAATCAAACTCTGAGAGTCAATATCTGTACC 6767
Db 793 nHisLeuCysAsnHisThrGluAlaPheAenArgThrSerGluSerGlnTyrrLeuValPr 813
QY 6768 TGCTTATGACAGGTGGGCATGTAATATACGATTAAACCCCTTGTTCCACCTTGGT 6827
Db 813 oGlyTyrrAspArgTrpAlaCysaenThrGlyLeuThrProCysValSerThrLeuVa 833
QY 6828 TTTTAAACCAACTAAAGATTTTGCATTATGTGTCCTCAAAATGTTCCCGAGTGTATTACTA 6887
Db 833 lPheAenGlnThrLysAspPheCysValMetValGlnIleValProArgValTyrrTy 853
QY 6888 TCCGAAAAGCAATCTCTGATGAATATGACTACAGAAATCATCGCAAAAGAGAGAACC 6947
Db 853 rProGluLysAlaValLeuAspGluTyrrAspTyrrArgTyrrAsnArgProLysArgGluPr 873
QY 6948 CATATCTCTGACACTTCTGCTGTCCTCGACTTGGAGTGGCAGCAGGTGTAGGAACAGG 7007
Db 873 oIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThrGI 893
QY 7008 AACAGTCTCCCTGCTGACCGGACACACAGCAGCTAGAAACAGGACTTAGTAACCTACATCG 7067
Db 893 yThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSerAsnLeuHisAr 913
QY 7068 AATTGTAACAGAGATCTCAAGCCCTAGAAAATCTGTCACTGAGGAGAACTCCCT 7127
Db 913 gIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSerLe 933
QY 7128 AACCTCTTATCTGAAGTACTCTACAGATAGAGAGGCTTAGATTATTCTTAA 7187
Db 933 uThrSerLeuSerGluValValLeuGlnAenArgArgGlyLeuAspLeuLeuPheLeuLy 953
QY 7188 AGAAGAGAGATTATGTAGCTTTGAAGAGGAATGCTGTTTTATGTGATCATTCAGG 7247
Db 953 sGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrrValAspHisSerGI 973
QY 7248 GGCATCAGAGATCTCCATGAACAGCTTAGAGAAAGTTGGAGAGCGTCGAAGGGAATA 7307
Db 973 yAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArgArgGluAr 993
QY 7308 GGAACCTACTCAAGGTGGTTTCAGGATGCTTCACAGCTCTCTTGTGGTACCCCT 7367
Db 993 gGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThrLe 1013
QY 7368 ACTTTCCTGTTTAAACAGGACCTTAAATAGTCCCTCTCTCTTACTCACAGTTGGGCCATG 7427
Db 1013 uLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThrValGlyProCy 1033
QY 7428 TATTATTAACAAGTTAATTCCTTCAATTAGAGAACGAATAGTCAGTCCAGATCATGGT 7487
Db 1033 sLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaValGlnIleMetVa 1053
QY 7488 ACTTAGACACAGTACCAG 7506
Db 1053 lLeuArgGlnGlnTyrrGln 1059

RESULT 9

ID AAW81572 standard; protein; 1203 AA.
XX AC AAW81572;
XX XX
DT 01-MAR-1999 (first entry)
XX Mus dunni endogenous virus Pol protein.
DE MDEV; retrovirus; packaging cell line; gene transfer; gene therapy;
KW vector; polymerase protein; Pol protein.
XX XX
OS Mus dunni endogenous virus.

Alignment Scores:
Pred. No.: 0 Length: 1203
Score: 4332.50 Matches: 822
Percent Similarity: 79.6% Conservative: 129
Best Local Similarity: 68.8% Mismatches: 222
Query Match: 29.6% Indels: 21
DB: 2 Gaps: 7

US-10-723-552-3 (1-8132) x AAW81572 (1-1203)

QY 2160 GGGAGAGCGGGTTTCGGACCCCTCCCGAGCCCGAGGTAACCTTTGAAGGTGGAGGGCAA 2219
Db 1 GlySerArgGlySerAspProLeuProGluProArgValThrLeuSerValGluGlyThr 20
QY 2220 CCAGTTGAGTTCCTGCTGTGATACCGAGCGAATTCAGTGTCTACTACAGCCATTAGGA 2279
Db 21 ProValAsnPheLeuIleAspThrGlyAlaGluHisSerValLeuThrSerProLeuGly 40
QY 2280 AAACATAAAGATAAAAAATCCTCGGTGATGGTGGCCACAGGGCAACACAGTATCCATGG 2339

XX WO9850538-A1.
PN 12-NOV-1998.
XX 08-MAY-1998; 98WO-US009452.
PD 09-MAY-1997; 97US-0046140P.
XX 08-MAY-1998; 98US-00075272.
PF (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX Miller AD, Wolgamot G, Bonham L;
XX WPI; 1999-034718/03.
DR N-PSDB; AAV69750.
XX New retroviral packaging cells - containing Mus dunni endogenous virus
PT sequences to target cells, retrovirus gag and pol genes and a
PT heterologous gene of interest.
XX Disclosure; Page 68-71; 85pp; English.
XX This is the amino acid sequence of the polymerase (Pol) protein of Mus
CC dunni endogenous virus (MDEV), as deduced from the MDEV nucleotide
CC sequence (see AAV69750). A cultured packaging cell is claimed which
CC produces a replication-defective retroviral vector (RDRV) particle, where
CC the packaging cell is a vertebrate cell capable of expressing and
CC assembling retroviral proteins, comprising: (a) a first vector encoding a
CC retroviral envelope protein having amino acid residues MDEV that direct
CC binding of the retroviral particle to MDEV retroviral receptors on a
CC target cell; and (b) a second vector encoding retrovirus gag and Pol
CC proteins, where upon expression of the vectors in the packaging cell in
CC the presence of a vector having a sequence of a heterologous gene of
CC interest, a replication-defective retroviral particle is produced that
CC binds to MDEV receptors of target cells. Also claimed are: (1) a cultured
CC packaging cell for producing a RDRV particle; (2) methods for producing a
CC RDRV particle comprising a heterologous gene of interest; (3) cultured
CC packaging cell line PD223; and (4) a RDRV produced by a method as in (2).
CC The MDEV receptor is present on a variety of cells rendering MDEV
CC pseudotype packaging cells useful in methods of mammalian and
CC particularly human gene transfer for gene therapy. The MDEV packaging
CC cells are a stable and reproducible source of retroviral particles.
CC Clones may be isolated from these populations that produce high titre
CC virus. The packaging cell lines may be selected and cloned for other
CC desirable properties, such as stability of in vivo growth, lack of
CC production of helper virus, lack of reinfection by viral particles
CC packaged in the cell, stability from genetic rearrangement and
CC recombinational events, resistance to complement lysis, and improved
CC ability to infect cells from higher mammals

SQ Sequence 1203 AA;

41	LysLeuGlySerLysArgThrIleValValGlyAlaThrGlySerLysLeuTyrProTyr	60
2340	ACTACCCGAAGAACAGATTGACTTGGAGTGGAGCGGGTAACCCACTCGTTTCTGCTCAT	2399
	: : : : :	
61	ThrThrLysArgAlaLeuGlnIleAspLysAenMetValThrHisSerPheLeuValIle	80
2400	CCTGAGTCCCGACGACCCCTCTTAGGTAGACACTATTGACCAAGATGGGAGCGCAAAAT	2459
	: : : : :	
81	ProGluCysProAlaProLeuLeuGlyArgAspLeuLeuThrLysLeuLysAlaGlnVal	100
	: : : : :	
2460	TCTTTTGAACAGGGAACAGAGTGTCTCAATAACAAACCTATCATCTGCTGTTCACC	2519
	: : : : :	
101	GlnPheThrSerGluGlyProGlnValSerTrpGlyLysAlaProLeuAlaCysLeuVal	120
	: : : : :	
2520	CTCCAATTAGATGACGAATATCGACTATACTCTCTCCCTAGTAAAGCCGTGATCAAAAT	2579
	: : : : :	
121	LeuSerThrGluGluGluTyrArgLeuHisGluGluGlnProLysGlyAlaAlaProLeu	140
	: : : : :	
2580	CAATTCGTGGTGGAACTGTTTCCCCAGCCTGGGCGAGAAACCGCAGGGATGGGTTGGCA	2639
	: : : : :	
141	Asp---TrpValThrAlaPheProAsnValTrpAlaGluGlnAlaGlyMetGlyLeuAla	159
	: : : : :	
2640	AAGCAAGTTCCTCCCAAGTATTCACTCGAAGCCAGTCGCCACACAGTGTCTAGTCAGA	2699
	: : : : :	
160	LysGlnValProProValValValGluLeuLysAlaAspAlaThrProLysValArg	179
	: : : : :	
2700	CAGTACCCCTTGACTAAAGAAGCTCAAGAGAGAAATTCGGCGCATGTCCAAAGATTATC	2759
	: : : : :	
180	GlnTyrProMetSerLysGluAlaLysGluGlyLeuArgProHisIleArgArgLeuLeu	199
	: : : : :	
2760	CAACAGGGCATCTAGTTCCTGTCCTCAATCTCCCTGGAAATCTCCCTGTCTACCGGTTAGA	2819
	: : : : :	
200	AspGlnGlyIleLeuValAlaCysGlnSerProTrpAsnThrProLeuLeuProValArg	219
	: : : : :	
2820	AAGCCTGGCATATGACTATCGACAGTACAGCACTTGAGAGAGGTCAATAAACGGTG	2879
	: : : : :	
220	LysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgVal	239
	: : : : :	
2880	CAGGATATACACCCAAACAGTCCCGAACCTTATAACCTCTGTGTGCTCTCCACCCCAA	2939
	: : : : :	
240	LeuAspIleHisProThrValProAsnProTyrAsnLeuLeuSerSerLeuProGlu	259
	: : : : :	
2940	CGGAGCTGGTATACAGTATTGGACTTAAGAGATGCCCTTCTCTGCTGAGATTACACCCC	2999
	: : : : :	
260	ArgThrTrpTyrThrValLeuAspLeuLysAspAlaPheCysLeuArgLeuHisPro	279
	: : : : :	
3000	ACTAGCCAAACCACTTTTGGCTTCGATGAGAGATCCAGGTACGGGAGAACCGGGCAG	3059
	: : : : :	
280	LysSerGlnLeuLeuPheAlaPheGluTrpArgAspProGluGlyGlyGlnThrGlyGln	299
	: : : : :	
3060	CTACCTGGAGCCGACCTGCCCAAGGTTCAAGAACTCCCGACCATCTTTGACGAGCC	3119
	: : : : :	
300	LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAspGluAla	319
	: : : : :	
3120	CTACACAGAGACCTGGCCAACTTCAGGATCCAAACACCTCAGGTGACCTCTCCAGTAC	3179
	: : : : :	
320	LeuHisArgAspLeuAlaProPheArgAlaGlnAsnProGlnLeuThrLeuLeuGlnTyr	339
	: : : : :	
3180	GTGATGACCTGCTCTGGCGGGAGGCCAACAAACAGGACTGCTTAGAAGGACGAAAGGCA	3239
	: : : : :	
340	ValAspAspLeuLeuIleAlaAlaAspLysGluLeuCysGlnGlnGlyThrGluArg	359
	: : : : :	
3240	CTACTGTGGAATTGCTGACCTAGGCTACAGAGCCCTGCTAAGNAGGCCCGCATTTGC	3299
	: : : : :	
360	LeuLeuThrGluLeuGlyAsnLeuGlyTyrArgValSerAlaLysLysAlaGlnIleCys	379
	: : : : :	
3300	AGGAGAGAGGTAAATCTTTGGGGTACAGTTTGGGACCGGCGCAGCATGGCTGACCGAG	3359
	: : : : :	
380	GlnThrGluValIleTyrLeuGlyTyrThrLeuArgGlyLysArgTrpLeuThrGlu	399
	: : : : :	
3360	GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACCAACGCAAAACAAATGAGAGAG	3419
	: : : : :	
400	AlaArgLysLysThrValMetMetIleProProThrThrProArgGlnValArgGlu	419
	: : : : :	

Qy	3420	TTTTTGGGGA	CAGCTGGATTTTGCAGACTGTGGATCCCGGGTTTGCAGCTTAGCAGCC	3479
Db	420	PhleuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAla	439	
Qy	3480	CCACTTACCCGCTAA	CCAAAGAAAAGGGGAATTCCTCTGGGCTCCTGAGCACCAGAAG	3539
Db	440	ProLeuTyProLeuThrArgGluGlyIleProPheGluTrpIysGluGluHisGlnArg	459	
Qy	3540	GCATTTGATGCTATCA	AAAAAGCCCTGCTCAGGCGCACCTGCTCTGGCGCCCTCCCTGACGTA	3599
Db	460	AlaPheGluAlaIleLysSerSerLeuMetThrAlaProAlaLeuAlaLeuProAspLeu	479	
Qy	3600	ACTAAACCTTTACCT	TTATGTGGATGAGCGTAAGGAGTAGCCGGGAGTTTAAACC	3659
Db	480	ThrLysSerPheValLeuTyrValAspGluArgAlaGlyIleAlaArgGlyValLeuThr	499	
Qy	3660	CAAAACCTTAGACCAT	GAGAGACCTGTCGCCTACCTGTCAAGAAGCTCGATCCTGTA	3719
Db	500	GlnAlaLeuGlyProTrpLysArgProValAlaTyrlLeuSerLysLysLeuAspProVal	519	
Qy	3720	GCCAGTGGTGGCCCAT	ATGCTGAAGGCTATCGCAGCTGTGCGCATCTGCTCAAGGAC	3779
Db	520	AlaSerGlyTrpProThrCysLeuLysAlaIleAlaValaLeuLeuIleLysAsp	539	
Qy	3780	GCTGACAAATTGACT	TTTGGGACAGAAATTAACCTGTAATAGCCCCCATGCAATTCGGAGAAC	3839
Db	540	AlaAspLysLeuThrMetGlyGlnGlnValThrValValaAlaProHisAlaLeuGluSer	559	
Qy	3840	ATCCTTGGCAGCCCC	CAGACCGATGATGACCAACGCCCGCATGACCCACTATCAAAAGC	3899
Db	560	IleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrlGlnSer	579	
Qy	3900	CTGCTCTTCACAGAG	GGGTACGTTCCGCTCCACCGCCGCTCTCAACCCCTGCCACTCTT	3959
Db	580	LeuLeuLeuAsnArgValThrPheAlaProProAlaIleLeuAsnProAlaThrLeu	599	
Qy	3960	CTGCCTGAAGAGACT	GTAGTAACCACTGACTCATGATTGCCATCAACTATTGATTGAGGAG	4019
Db	600	LeuProLeuThrAsnAspSerValProValHisArgCysAlaAspIleLeuAlaGluGlu	619	
Qy	4020	ACTGGGTCCGCAAG	CACTTACAGACATACCGCTGACTGGAGAAGTGCTAACTGGTTC	4079
Db	620	IleGlyThrArgLysAspLeuThrAspGlnProTrpTrpProGly---AlaProSerTrpTyrl	638	
Qy	4080	ACTGACGGAAGCAG	CTATGTGTGGAGGTAAAGAGATGCTGGCGCGCGGTGGTGAC	4139
Db	639	ThrAspGlySerPheLeuIleGluGlyLysArgArgAlaGlyAlaAlaValAlaAsp	658	
Qy	4140	GGGACCCGACAGAT	CTGGGCGCAGCGCTCCCGGAAGGAACTTTCAGCACAAAAGGCTGAG	4199
Db	659	GlyLysLysValIleTrpAlaSerAlaLeuProGluGlyThrSerAlaGlnLysAlaGlu	678	
Qy	4200	CTCATGGCCCTCAG	CAAGCTTTTGGCGCTGGCCGAAGGGAAATTCATAAACATTTATACG	4259
Db	679	LeuIleAlaLeuThrGlnAlaLeuArgGluAlaGluGlyIleAlaSerIleTyrlThr	698	
Qy	4260	GACAGCAGGTATGC	CTTTGGCAGCTGCACACGTCATGTTGGGCGCATCTATATAACAAAGGGGG	4319
Db	699	AspSerArgTyrlAlaPheAlaThrAlaHisIleHisGlyAlaIleTyrlArgGlnArgGly	718	
Qy	4320	TTGCTTACCTCAG	CGGGGGAATTAAGAACAAAGAGGAATTCCTCAACCTATTAGAA	4379
Db	719	LeuLeuThrSerAlaGlyLysAspIleLysAsnLysGluIleLeuAlaLeuLeuGlu	738	
Qy	4380	GCAGTCATTTACA	AAAAAGCGTAGCTATTATACACTGCTCTGGAATCAGAAAGCTAA	4439
Db	739	AlaIleHisAlaProLysLysValAlaIleIleHisCysProGlyHisGlnLysGlyGlu	758	
Qy	4440	GATCTCATATCC	AGAGAAACAGATGGCTGACCGGTTGCCAAGCAGGCGACCCAGGGT	4499
Db	759	AspLeuValAlaLysGlyAsnArgMetAlaAspSerValAlaLysGlnValaLysGlnGly	778	

```
QY 4500 GTTAACTCTCTGCTATATAGAAATGCCCC--AAAGCCCCAGAAACCCAGACGACAGTAC 4556
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
779 AlaMetIleLeuThrGluLysGlyAsnProSerLysSerProGlu-----793
QY 4557 ACCCTAGAAGACTGGCAAGAGATATAAAAGATAGACAGTTCCTGAGACTCCGGAAGGG 4616
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
794 -----AepGluAenTyrAspIleLysGluLeuPheTrpThrSerAspPro 808
QY 4617 ACCTGCTATACCTCAGATGGGAAGGAATCTCTCCCCACAAAGAGGGTTAGATATGTC 4676
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
809 LeuProTyrPhePheGluGlyLysIleAepLeuThrProGluGluGlyIleLysPheVal 828
QY 4677 CAACAGATACATCGTCTAACCCACCTAGGAATCTAAACACCTGCAGCAGTGTGGTCAGAAC 4736
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
829 LysGlyLeuHisGlnPheThrHisLeuGlyValGluLysMetMetArgLeuIleLysLys 848
QY 4737 TCCCTTATCATCTTCGAGGCTACCAGGAGTGGCTGACTCGGTGTGTCAAACATTTGTGTG 4796
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
849 SerArgTyrGlnValProAenLeuLysSerValAlaGlnLysIleIleAenSerCysLys 868
QY 4797 CCTGTCCAGCTGTTATGCTAATCTCTCCAGATGCTCTCCAGGAAGAGACTAAGGGGA 4856
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
869 AlaCysAlaPheThrAenAlaThrLysThrLysGluProGlyLysArgGlnAArgLys 888
QY 4857 AGCCACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGCTAAATACGGA 4916
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
889 AspArgProGlyValTyrTrpGluValAspPheThrGluValLysProGlyMetTyrGly 908
QY 4917 AACAAATACCTATGTTTTTTAGACACCTTTTTCAGGATGGGTAGAGCTTATCTACT 4976
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
909 AsnLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaPheProThr 928
QY 4977 AAGAAAGAGACTTCAACCGTGGTGGCTTAAATAATCTGCAAGAAATTTTCCAAATTT 5036
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
929 LysThrGluThrAlaGlnIleValAlaLysLysIlePheGluGluIleLeuProArgTyr 948
QY 5037 GGAATACCTAAGTAAATAGGTCAGACAATGGTCCAGCTTTTGTGCCAGGTAAGTCAG 5096
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
949 GlyValProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGln 968
QY 5097 GGACTGGCCCAAGATATTGGGGATTGATTGGAATCTGCATTTGTCATACAGACCCCAAGC 5156
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
969 GlyLeuAlaThrGlnLeuGlyIleAepTrpLysLeuHisCysAlaTyrArgProGlnSer 988
QY 5157 TCAGGCAGGTAGAGAGTGAATAGAACCTTAAAGAGACCTTACTAATTAATGACCGCG 5216
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
989 SerGlyGlnValGluArgMetAenArgThrLeuLysGluThrLeuThrLysLeuAlaMet 1008
QY 5217 GAGACTGGCGTTAATGATTGGATAGCTCTCCTGCCCTTTGTGCTTTTGTAGGTTAGGAAC 5276
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1009 GluThrGlyLysAspTrpValAlaLeuLeuProLeuAlaLeuPheArgAlaAArgAsn 1028
QY 5277 ACCCTGGACAGTTGGCTGACCCCTATGAATTACTTACGGGGAGACCCCCCATTTG 5336
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1029 ThrProGlyArgPheGlyLeuThrProPheGluValLeuTyrGlyGlyProProLeu 1048
QY 5337 GTAGAAATCTCTCT-----GTACATAGTGTGACGTGCTGCTTTCCCGACCTTTGTTTC 5390
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1049 IleLysAepGlyGlyThrLeuValProAepSerGlySerValLeuProSerSerLeuLeu 1068
QY 5391 TCTAGGCTCAAGCACCTTGAAGTGGTGAGACAACAGCAGCGTGGAGCAACTCCGGGAGGCC 5450
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1069 IleHisLeuLysAlaLeuLysValIleArgThrGlnIleTrpAspGlnLeuLysThrAla 1088
QY 5451 TACTCAGGAGGAGAGACTTGCAGATCCCATCATCGTTTCCAAAGTGGAGATTCAGTCTAC 5510
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1089 TyrThr---ProGlyThrThrAlaValProHisGluPheGlnValGlyAepGlnValLeu 1107
QY 5511 GTTAGAGCCACCGTCAGCAAAACCTCGACACTCGGTGGAGGCGCTTATCTCGACTT 5570
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1108 ValArgArgHisArgThrGlySerLeuGluProArgTrpLysGlyProTyrLeuValLeu 1127
QY 5571 TTGACCACACCAACGGCTGTGAAAGTCGAAGGAATCTCCACCTGGATCCATGATCCAC 5630
```

```
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1128 LeuThrThrProThrAlaValLysValAspGlyIleAlaSerTrpIleHisAlaSerHis 1147
QY 5631 GTTAAACCGCGCCACCTCCGAT-----TCGGGGTGGAAACCGGAAAG 5675
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1148 ValLysArgAlaProSerGlnAspGluThrHisGluAspAsnTrpAlaValGluAla 1167
QY 5676 ACTGAAATCCCTTAAGCTTCGCCCTCCATCGCGTGGTTCCCT 5717
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1168 ThrAspAsnProLeuLysLeuArgLeuArgArgSerPro 1181

RESULT 10
AAR75189
ID AAR75189 standard; protein; 1196 AA.
XX
AC AAR75189;
XX
DT 27-AUG-2003 (revised)
DT 30-MAY-1996 (first entry)
XX
DE Osteoinductive retrovirus RFB-14 pol gene product.
XX
KW RFB retrovirus; gag; pol; env; osteogenesis; osteoinductive protein;
KW bone development; osteoporosis; gene therapy; polymerase;
KW reverse transcriptase.
XX
OS Murine leukemia virus.
XX
PN DE4411718-Al.
XX
PD 12-OCT-1995.
XX
PF 05-APR-1994; 94DE-04411718.
XX
PR 05-APR-1994; 94DE-04411718.
XX
PA (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEIT.
XX
PI Schmidt J, Gimbel W, Strauss P, Erfle V, Pedersen PS, Pedersen L;
PI Oestergaard M;
XX
DR WPI; 1995-352078/46.
DR N-PSDB; AAQ94266.
XX
PT RFB-14 retrovirus genome - and prodn. of osteo-inductive proteins.
XX
PS Claim 14; Fig 1; 46pp; German.
XX
CC The full-length proviral genomic sequence of retrovirus RFB-14 has been
CC determined. The virus codes for an osteoinductive protein, although the
CC precise location of the coding region has not yet been identified. The
CC virus may be useful in gene therapy of bone growth disorders such as
CC osteoporosis. The present sequence is that of the viral pol gene product.
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 1196 AA;

Alignment Scores:
Pred. No.: 2,66e-294 Length: 1196
Score: 3981.00 Matches: 752
Percent Similarity: 76.8% Conservative: 168
Best Local Similarity: 62.8% Mismatches: 264
Query Match: 27.2% Indels: 14
DB: Gaps: 9

US-10-723-552-3 (1-8132) x AAR75189 (1-1196)
QY 2160 GGGAGACGGGTTTCGACCCCTCCCGAGCCAGCGTAACCTTGAAGGTGAGGGCGCA 2219
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1 GlyGlyGlnGlyGlnGluProProGluProArgIleThrLeuThrValGlyGln 20
QY 2220 CCAGTTGAGTTCCTGTTGATACCGGAGCGAAACATTTCAGTGTACTACAGCCATTAGGA 2279
Db      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Db 21 ProValThrPheLeuValAspThrGlyAlaGlnHisSerValLeuThrGlnAsnProGly 40
QY 2280 AAACTAAAGATAAAATAATCTGGGTGATGGGTGCCACAGGGCAACAACAGTATCCATGG 2339
Db 41 ProLeuSerAspArgSerAlaTvpValGlnGlyAlaThrGlyGlyLysArgTyrArgTtp 60
QY 2340 ACTACCGAAGAACAGTGTGACTGGGAGTGGGAGCGGTAAACCACTCGTTCTTCGGTCATYA 2399
Db 61 ThrThrAspArgLysValHisLeuAlaThrGlyLysValThrHisSerPheLeuHisVal 80
QY 2400 CTGTAGTGGCCAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAAT 2459
Db 81 ProAspCysProTyrProLeuLeuGlyArgAspLeuLeuThrLysLeuLysValGlnIle 100
QY 2460 TCTTTTGA---CAAGGAAACAGAGTGTCTGCAAAATAACAACCTATCACTGTGTG 2516
Db 101 HisPheGluGlySerGlyAlaGlnValValGlyProLysGlyGlnProLeuGlnValLeu 120
QY 2517 ACCCTCAATTAGATGACGATATGACTATATCTCTCCCTAGTAAAGCTGATCAAAAT 2576
Db 121 ThrLeuAsnLeuGluAspGluTyrArgLeuTyrGluThrSerAlaGluProGluAlaSer 140
QY 2577 ATA---CAATTCTGTTGGAACAGTTTCCCAAGCTGGGCAGAAACCGCAGGGATGGT 2633
Db 141 ProGlySerThrTrrPheSerAspPheProGlnAlaTrrPalaGluThrGlyMetGly 160
QY 2634 TTGGCAAGCAAGTCTCCCAACAAAGTTATTCAACTGAAGCCAGTCCACACAGTGTCA 2693
Db 161 LeuAlaValArgGlnAlaProLeuIleIleProLeuLysAlaThrSerThrProValSer 180
QY 2694 GTCAGACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCGGAGTGTCCAAAGA 2753
Db 181 IleLysGlnTyrProMetSerGlnGluAlaArgLeuGlyIleLysProHisIleGlnArg 200
QY 2754 TTAATCAACAGGGCATCTAGTTCCTGTCCTGATCCCTGGAATCTCCCTGCTACCG 2813
Db 201 LeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrrPasnThrProLeuLeuPro 220
QY 2814 GTTAGAAGCCTGGGACTAATGACTATCGACCACTACAGGACTTGAGAGAGGTCAATAAA 2873
Db 221 ValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLys 240
QY 2874 CGGTGCAGGATATACCCCAACAGTCCGGAACCTTTATAACCTCTTGTGTCTCTCCCA 2933
Db 241 ArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeuPro 260
QY 2934 CCCCAACGGAGCTGSPATACAGTATTGGACTTAAAGGATCCCTCTTCTGCTGAGATTA 2993
Db 261 ProSerHisArgTrrPyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeu 280
QY 2994 CACCCCACTAGCCCAACCACTTTTTCGCTTCGAATGGAGAGATCCAGGTACGGGAAGAAC 3053
Db 281 HisProThrSerGlnProLeuPheAlaPheGluTrrPArgAspProGlyMetGlyIleSer 300
QY 3054 GGGCAGCTCACTGGACCCGACTGCCCAAGGTTCAAGACTTCCCGACCACTTTTGAC 3113
Db 301 GlyGlnLeuThrTrrPThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAsp 320
QY 3114 GAAGCCCTACAGAGACTGGCCAACTTCAGATCCCAACCCCTCAGGTGACCTCTCTC 3173
Db 321 GluAlaLeuHisArgAspLeuAlaGlyPheArgIleGlnHisProAspLeuIleLeuLeu 340
QY 3174 CAGTACGTGGATGACCTGCTTCTGGCGGGAGGCCAACCAACAGGACTTCGTAGAAGGCAG 3233
Db 341 GlnTyrValAspAspLeuLeuLeuAlaAlaThrSerGluLeuAspCysGlnGlnGlyThr 360
QY 3234 AAGGCATCTACTGCTGGAATTGTCTGACTAGGCTACAGAGCTCTGCTTAAGAGGCCAC 3293
Db 361 ArgAlaLeuLeuGlnThrLeuGlyAspLeuGlyTyrArgAlaSerAlaLysLysAlaGln 380
QY 3294 ATTTGCAGAGAGAGTACATCTTGGGGTACAGTTTCGGGAGCGGAGGAGGTGGCTG 3353
Db 381 IleCysGlnLysGlnValLysTyrLeuGlyTyrLeuLeuLysGluGlyGlnArgTrrPLeu 400

QY 3354 ACGGAGCCAGCGAAGAAATACTGTAGTCCAGATACCGGCCCCCAACACACAGCAAAATG 3413
Db 401 ThrGluAlaArgLysGluThrValMetGlyGlnProIleProLysThrProArgGlnLeu 420
QY 3414 AGAGAGTTTTTGGGACACCTGGATTTCAGACTGTGAGTCCCGGGTTTGGACCTTA 3473
Db 421 ArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrrPileProGlyPheAlaGluMet 440
QY 3474 GCAGCCCACTCTACCCGCTAAACCAAGAAAGGGAAATTCCTCTGGGCTCTGAGCAC 3533
Db 441 AlaAlaProLeuTrrPrrProLeuThrLysThrGlyThrLeuPheAsnTrrPrrGlyProAspGln 460
QY 3534 CAGAAGGCATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTCTCTGGGCCCTCCCT 3593
Db 461 GlnLysAlaTyrGlnGluIleLysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeuPro 480
QY 3594 GAGCTAACTAAACCTTTTACCTTTATGTGTGATGACGTAAAGGATAGCCCGGGAGTT 3653
Db 481 AspLeuThrLysProPheGluLeuPheValAspGluLysGlnGlyTyrAlaLysGlyVal 500
QY 3654 TTAAACCAACCTAGGACCATGGAGAACCTGTGCTGCTACCTGTCAAGAGACTCGAT 3713
Db 501 LeuThrGlnLysLeuGlyProTrrPArgSerValAlaTrrLysSerLysLysLeuAsp 520
QY 3714 CCTGTAGCCAGTGGTGGCCCATATGCTCTGAAGGCTATCCAGCTGTGGCCATCTGGTC 3773
Db 521 ProValAlaAlaGlyTrrPrrProCysLeuArgMetValAlaAlaIleAlaValLeuThr 540
QY 3774 AAGGACCTGACAAATTTGATTTGGGACAGATATTAATGTATATAGCCCCCATGATTG 3833
Db 541 LysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAlaGlu 560
QY 3834 GAGAACATCTGTCGGCAGCCCGACGATCGATGACCAACGCGCATGACCCACTAT 3893
Db 561 GluAlaLeuValLysGlnProAspArgTrrPLeuSerAsnAlaAlaArgMetThrHisTyr 580
QY 3894 CAAAGCTGCTTCTC---ACAGAGAGGTTCAGTTCGCTCCACCCAGCCGCTCTCAACCCCT 3950
Db 581 GlnAlaMetLeuLeuAspThrAspArgValGlnPheGlyProValValAlaLeuAsnPro 600
QY 3951 GCCACTTCTGCTGAGAGACTGATGAACACGATGACTCATGATGTTGCCATCACTATTG 4010
Db 601 AlaThrLeuLeuPro---LeuProGluGluGlyAlaProHisAspCysLeuGluIleLeu 619
QY 4011 ATTGAGGAGACTGGGTCCGCAAGGACCTTACAGACATACCGCTGACTCGAAGTGCCTA 4070
Db 620 AlaGluThrHisGlyThrArgProAspLeuThrAspGlnProIleProAspAlaAspHis 639
QY 4071 ACCTGTTCTACGAGGAGCAGCTATGTGTGGAAGGTAAAGAGATGCTGGCGCGCG 4130
Db 640 ThrTrrPrrSerAspGlySerSerPheLeuGlnGluGlyGlnArgLysAlaGlyAlaAla 659
QY 4131 GTGTGGAGCGGACCGCACGATCTGGGCCAGACGCTCCCGAAGGAATCTTCAGACAA 4190
Db 660 ValThrThrGluThrGluValIleTrrPAlaArgAlaLeuProAlaGlyThrSerAlaGln 679
QY 4191 AAGGCTGAGCTCATGCGCTCACCAAGCTTTGCGGCTGGCGAAGGGAATCCATAAC 4250
Db 680 ArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysArgLeuAsn 699
QY 4251 ATTTATACGAGCAGCAGGTATGCTTTTGCAGCTGCACACGCTACATGGGGCCATCTATAA 4310
Db 700 ValTrrThrAspSerArgTrrAlaPheAlaThrAlaHisIleHisGlyGluIleTrrArg 719
QY 4311 CAAAGGGGTGCTTACTCTACAGAGGAGGAAATAAAGAAACAAGAGAAATCTTAACG 4370
Db 720 ArgArgGlyLeuLeuThrSerGluGlyArgGluIleLysAsnLysSerGluIleLeuAla 739
QY 4371 CTATTAGAGCGGTACATTTTACCAAAAGGCTAGCTATTATATACATGCTCTGACATCAG 4430
Db 740 LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysLeuGlyHisGln 759

QY	4431	AAAGCTAAAGATCTCATATCCAGAGGAAACCCAGATGGCTGACCGGGTTGGCCAGCAGGCA	4499
Db	760	LyseGlyAspSerAlaGluAlaArgGlyAsnArgLeuAlaAspGlnAlaAlaArgGluAla	779
QY	4491	GCC-----CAGGGTGTTAACTTCTGCTATATAATAGAAATGCCCAAGGCCCA	4538
Db	780	AlaIleLysThrProProAspThrSerThrLeuLeuIleGluAspSerThrProTyThr	799
QY	4539	GAACCCAGACGACATAGACCTTACAGATGGCGACAGATATAAAAGATAGACCAAGTTC	4598
Db	800	ProAlaTyRpheHisTyThrGluThrAspLeuLysLysLeuArgGluLeuGlyAlaThr	819
QY	4599	TCTGAGACTCCGGAAGGACCTGCTATACCTCAGATCGGAAGAAATCTCGCCCAAAA	4658
Db	820	TyrAsnGlnSerLysGly--TyrTrpValPheGlnGlyLysProValMetProAspGln	838
QY	4659	GAAGGTTAGATATGTCACACAGATACATCGTCTAACCCACCTAGGAACATAACACCTG	4718
Db	839	PheValPheGluLeuLeuAspSerLeuHisArgLeuThrHisLeuGlyTyRglnLysMet	858
QY	4719	CAGCAGTTGGTC-----AGAACATCCCTTATCATGTCTTGAGGCTACCAAGAGTGGCT	4772
Db	859	LysAlaLeuLeuAspArgGlyGluSerProTyRThrMetLeuAsnArgAspLysThrLeu	878
QY	4773	GACTCGGTGTTCAACATTTGTGTCCCTGCCAGCTGTTAATGCTAATCTCCAGAATG	4832
Db	879	GlnTyRValAlaAspSerCysThrValCysAlaGlnValAsnAlaSerLysAlaLysIle	898
QY	4833	CCTCCAGGGAAGACATAAGGGGAAGCCACCAGGCGCTCAGTGGGAAGTGGACTTCACT	4892
Db	899	GlyAlaGlyValArgValArgGlyHisArgProGlySerHisTrpGluIleAspPheThr	918
QY	4893	GAGTAAAGCCGGCTAAATACGGAACAAATACCTATTGGTTTGTGAGACACCTTTTCA	4952
Db	919	GluValLysProGlyLeuTyRgTyRysTyRLeuLeuValPheValAspThrPheSer	938
QY	4953	GGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAATA	5012
Db	939	GlyTrpValGluAlaPheProThrLysArgGluThrAlaArgValValSerLysLysLeu	958
QY	5013	CTGGAAGAAATTTTCCAAGATTGGAAATACCTAAGGTAAATAGGGTCAGACAATGTCCTA	5072
Db	959	LeuGluGluIlePheProArgPheGlyMetProGlnValLeuGlySerAspAsnGlyPro	978
QY	5073	GCTTTTGTGCCAGGTAAAGTCAGGGACTGSCCAAGATATTGGGGATTGATGGAAACTG	5132
Db	979	AlaPheThrSerGlnValSerGlnSerValAlaAspLeuLeuGlyIleAspTrpLysLeu	998
QY	5133	CATTGTGCATACAGACCCCAAGCTCAGGACAGGTAGAGAGATGAATAGAACCATTA	5192
Db	999	HisCysAlaTyRArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLys	1018
QY	5193	GAGACCTTACTAAATTGACCGCGGAGACTGGCGCTTAATGATTGGATAGCTCTCTGCC	5252
Db	1019	GluThrLeuThrLysLeuThrLeuAlaAlaGlyThrArgAspTrpValLeuLeuLeuPro	1038
QY	5253	TTTGTGCTTTTAGGGTTAGAACACCCCTGGACAGTTTGGCTGACCCCTCATGAATTA	5312
Db	1039	LeuAlaLeuTyRArgAlaArgAsnThrProGlyProHisGlyLeuThrProTyRgIle	1058
QY	5313	CTCTACCGGGGACCCCCCATTCGTTAGAAATTGCTCTGTACATAGTGTGACGTGCTG	5372
Db	1059	LeuTyRgAlaProProProLeuValAsnPheHisAspProAspMetSerGluLeuThr	1078
QY	5373	CTTTCCAGCGCTTTGTTCTTAGGCTCAAGGCATCTGAGTGGGTGAGACAACGAGCGTGG	5432
Db	1079	AsnSerProSerLeuGlnAlaHisLeuGlnAlaLeuGlnThrValGlnArgGluIleTrp	1098
QY	5433	AGGCMACTCCGGGAGGCTACTCAGGAGGAGGACTTGCAG---ATCCCCACATCGTTTC	5489
Db	1099	LysProLeuAlaGluAlaTyRArgAspGlnLeuAspGlnProValIleProHisProPhe	1118
QY	5490	CAAGTGGGAGATTCACTACGTTTAGACGCCACCGTCAGAAACCTTCGAGACTCGGTGG	5549

Db	1119	ArgGlyLeuAspSerValTrpValArgHisGlnThrLysAsnLeuGluProArgTrp	1138
Qy	5550	AAGGGCCCTTATCTCGTACTTTTGACACACCAACGCGCTGTGAAGTCGAAGGAATCTCC	5609
Db	1139	LysGlyProIyrrThrValLeuLeuThrProThrAlaLeuLysValAspGlyIleSer	1158
Qy	5610	ACCTGGATCATGCATCCCATCGTTAAACCGCGC-----CCACCTCCCGATTCCGGGGTGG	5663
Db	1159	AlaTrpIleHisAlaAlaHisValLysAlaAlaThrThrProFroIleLysProSerTrp	1178
Qy	5664	AAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCT	5717
Db	1179	ArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGlyAlaPro	1196
RESULT 11			
ADH76470			
ID	ADH76470	standard; protein; 1199 AA.	
XX	AC	ADH76470;	
XX	DT	15-APR-2004 (first entry)	
XX	XX	Murine leukemia virus pol gene derived protein.	
DE	XX		
XX	XX	chimeric plasmid; replicative retroviral genome; gag; pol;	
KW	KW	murine leukemia virus; MLV; env; gibbon ape leukemia virus; GaLV; virion;	
KW	KW	MLV-GaLV-type; gene therapy; pAM.	
XX	OS	Murine leukemia virus.	
XX	XX		
FN	FN	FR2832424-A1.	
XX	XX	23-MAY-2003.	
PD	XX		
PF	PF	20-NOV-2001; 2001FR-00014976.	
XX	XX		
PR	PR	20-NOV-2001; 2001FR-00014976.	
XX	XX	(GENE-) GENETHON III.	
PA	XX		
XX	XX	Audit M, Cosset FL;	
FI	XX		
XX	XX	WPI; 2003-471779/45.	
DR	XX		
XX	XX		
PT	PT	Chimeric plasmid containing replicative retroviral genome, useful for	
FT	FT	making positive control virions in testing for replication-competent	
PT	XX	retrovirus.	
XX	XX		
PS	XX	Claim 4; SEQ ID NO 2; 70pp; French.	
XX	XX		
CC	CC	The invention relates to a novel chimeric plasmid comprising a	
CC	CC	replicative retroviral genome. The replicative retroviral genome	
CC	CC	comprises: the gag and pol sequences from a murine leukemia virus (MLV);	
CC	CC	and a chimeric env sequence comprising regions corresponding to parts of	
CC	CC	the envelope derived from: an MLV genome; and a gibbon ape leukemia virus	
CC	CC	(GaLV). Virions produced by expressing the viral genome of the chimeric	
CC	CC	plasmid are useful as positive controls in a test for detection of	
CC	CC	replication-competent retroviruses in preparations of MLV-GaLV-type	
CC	CC	retroviral vectors. For example, to ensure that the MLV-GaLV-type	
CC	CC	retroviral vectors, intended for gene therapy, have no capacity for	
CC	CC	replication. This sequence represents the protein derived from an MLV pol	
CC	CC	gene from a pAM plasmid of the invention.	
XX	XX		
SQ	SQ	Sequence 1199 AA;	
Alignment Scores:			
Pred. No.:	2,91e-294	Length:	1199
Score:	3980.50	Matches:	755
Percent Similarity:	76.2%	Conservative:	165
Best Local Similarity:	62.5%	Mismatches:	257
Query Match:	27.2%	Indels:	31
DB:	7	Gaps:	11

[illegible]

```
Db 660 ValThrThrGluThrGluValIleTrrAlaArgAlaLeuProAlaGlyThrSerAlaGln 679
QY 4191 AAGCTGAGCTCATGGCCCTCAGCAAGCTTGGGCTGCCGGAAGGAATCATAAAC 4250
Db 680 ArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysArgLeuAsn 699
QY 4251 ATTATACGACAGCAGGTATGCTTGGCAGCTGCACACTACATGGGCGCATCTATAA 4310
Db 700 ValTyrThrAspSerGlnTyrAlaPheAlaThrAlaHisIleHisGlyGluIleTyrArg 719
QY 4311 CAAAGGGGGTGTCTACTCAGCAGGAGGAGGAAATAAGAACAAAGAGGAAATCTTAAGC 4370
Db 720 ArgArgGlyLeuLeuThrSerGluGlyArgGluIleLysAsnLysSerGluIleLeuAla 739
QY 4371 CTATTAGAACCGTACATTATACCAAAAGCTAGCTATTATACATGCTCTGACATCAG 4430
Db 740 LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysLeuGlyHisGln 759
QY 4431 AAGCTAAAGATCTCATATCCAGAGGAACAGATGGCTGACCGGTGCGCAAGCAGGCA 4490
Db 760 LysGlyAspGlyAlaGluAlaArgGlyAsnArgLeuAlaAspGlnAlaAlaArgGluAla 779
QY 4491 GCC-----CAGGGTGTAACTTCTGCTTATAAGAAATGCCCAAGCCCCA 4538
Db 780 AlaIleLysThrProProAspThrSerThrLeuLeuIleGluAspSerThrProTyrThr 799
QY 4539 GAACCCAGACAGATACACCTAGAAAGATGCGCAAGATATAAAAGATAGACCATTC 4598
Db 800 ProAlaTyrPheHisTyrThrGluThrAspLeuLysLysLeuArgGluLeuGlyAlaThr 819
QY 4599 TCTGAGACTCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATCCTGCCCCACAA 4658
Db 820 TyrAsnGlnSerLysGly----TyrTrpValPheGlnGlyLysProValMetProAspGln 838
QY 4659 GAAGGGTGTAGATATGTCACACAGATACATCGTCTAACCCACCTAGGAATAAACACCTG 4718
Db 839 PheValPheGluLeuLeuAspSerLeuHisArgLeuThrHisLeuGlyTyrGlnLysMet 858
QY 4719 CAGCAGTTGTGTC-----AGAACATCCCTTATCATGTTCTGAGGCTACCCAGGAGTGCT 4772
Db 859 LysAlaLeuLeuAspArgGlyLysSerProTyrTyrMetLeuAsnArgAspLysThrLeu 878
QY 4773 GACTCGGTGTCAAACATCTGTGCTCCCTGCAGCTGGTAACTTAATCTTCCAGAAATG 4832
Db 879 GlnTyrValAlaAspSerCysThrValCysAlaGlnValAsnAlaSerLysAlaLysIle 898
QY 4833 CCTCCAGGGAAGACTAAAGGGGAAGCCACCCAGGCGCTCACTGGGAAGTGACTTCACCT 4892
Db 899 GlyAlaGlyValArgValArgGlyHisArgProGlySerHisTrpGluIleAspPheThr 918
QY 4893 GAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTTTTGTAGACACTTTTCA 4952
Db 919 GluValLysProGlyLeuTyrGlyTyrLysTyrLeuLeuValPheValAspThrPheSer 938
QY 4953 GGTGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTCAACCGTGGTGGCTAAATAATA 5012
Db 939 GlyTrpValGluAlaPheProThrLysArgGluThrAlaArgValValSerLysLysLeu 958
QY 5013 CTGAAGAAATTTTCCAGATTTGGAATACCTAAGTATAGGCTCAGACAAATGGTCCA 5072
Db 959 LeuGluGluIlePheProArgPheGlyMetProGlnValLeuGlySerAspAsnGlyPro 978
QY 5073 GCTTTTTCGCCAGGTAACTCAGGACTCGGCCAAGATATTTGGGATTTGATTTGAAACTG 5132
Db 979 AlaPheThrSerGlnValSerGlnSerValAlaAspLeuLeuGlyIleAspTrpLysLeu 998
QY 5133 CATTTGTCATACAGCCCCCAAGCTCAGGACAGTACAGAGGATGAATACACATTATA 5192
Db 999 HisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLys 1018
QY 5193 GAGACCCCTTACTAAATTTGACC---GCGGAGACTGGCGTTAATGATTTGATAGCTCTCTG 5249
```

```
Db 1019 GluThrLeuThrLysLeuThrLeuAlaAlaProGlyThrArgAspTrpValLeuLeuLeu 1038
QY 5250 CCCTTTGCTGCTTTTATAGGCTTAGACACACCCCTGGACAGCTTTTGGGCTGACCCCTATGAA 5309
Db 1039 ProLeuAlaLeuTyrArgAlaArgAsnThrProGlyProHisGlyLeuThrProTyrGlu 1058
QY 5310 TTACTCTACGGGGACCCCCCATTTGGTAGAAATTTGCTTCTGTACATAGCTGTCAGCTG 5369
Db 1059 IleLeuTyrArgAlaProProProLeuValAsnPheHisAspProAspMetSerGluLeu 1078
QY 5370 CTGCTTTCCAGCCTTTTGTCTCTAGGCTCAAGCACTTAGAGTGGGTGAGACAAACGAGCG 5429
Db 1079 ThrAsnSerProSerLeuGlnAlaHisLeuGlnAlaLeuGlnThrValGlnArgGluIle 1098
QY 5430 TGGAGCAACTCCGGGAGGCTTACTCAGGAGGAGGAGACTTGCGAG---ATCCACATCGT 5486
Db 1099 TrpAsnProLeuAlaGluAlaTyrArgAspGlnLeuAspGlnProValIleProHisPro 1118
QY 5487 TTCCAAGTGGGAGATTTCAGTCTACGTTAGACGCGCACGCTGCAGGAAACCTCGAGACTCGG 5546
Db 1119 PheArgIleGlyAspSerValTrpValArgHisGlnThrLysAsnLeuGluProArg 1138
QY 5547 TGAAGCGGCTTATCTCTGCTACTTTTGACACACCAACCGCTGTGAAAGTCGAAGAATC 5606
Db 1139 TrpLysGlyProTyrThrValLeuLeuThrThrProThrAlaLeuLysValAspGlyIle 1158
QY 5607 TCCACCTGGATCCATCATCCGACGCTTAACCGCGG-----CCACCTCCCGATTCCGGG 5660
Db 1159 SerAlaTrpIleHisAlaAlaHisValLysAlaAlaThrThrProProIleLysProSer 1178
QY 5661 TGGAAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCCTCCATCGCGTGTCTCT 5717
Db 1179 TrpArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGlyAlaPro 1197

RESULT 13
AAB12994
ID AAB12994 standard; protein; 1199 AA.
XX
AC AAB12994;
XX
DT 30-NOV-2000 (first entry)
XX
DE MLV reverse transcriptase protein sequence SEQ ID #41.
XX
KW Reverse transcriptase; RT; polymerase chain reaction; PCR; retrovirus;
XX stability; solubility.
XX Murine leukemia virus.
XX OS
XX
XX PN W0200042199-A1.
XX
XX PD 20-JUL-2000.
XX
XX PF 14-JAN-2000; 2000WO-US0000896.
XX
XX PR 15-JAN-1999; 99US-0116099P.
XX
XX PA (MOLE-) MOLECULAR BIOLOGY RESOURCES.
XX
XX Swaminathan N;
XX
XX WPI; 2000-482830/42.
XX
PT Novel genes encoding reverse transcriptase polypeptides modified by
PT altering or adding the integrase domains by truncation internally and/or
PT at the C-termini, useful in cDNA synthesis and amplification procedures.
XX
PS Claim 1; Page 154-159; 189pp; English.
XX
CC This invention relates to isolated polynucleotide sequences encoding a
CC polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT)
CC activity. RTs are found in a variety of retroviruses and their defining
CC activity is the ability to synthesise a cDNA strand using an RNA
```

CC template. The invention includes sequences AAA87808-A87822 and AAA87840-A87842 which represent nucleotide sequences encoding RT polypeptides. CC Included in the invention are reverse transcriptase protein sequences. CC AAB12989-B12995. Sequences AAA87829-A87839 and AAA87843-A87897 represent CC oligonucleotides used in the identification and synthesis of the RT CC nucleotide sequences of the invention. Many of the RT nucleotide CC sequences encode modified RT proteins, which exhibit improved stability CC and/or improved solubility, relative to naturally occurring reverse CC transcriptases. The modified RT nucleotide and protein sequence are used CC in improved polymerase chain reaction (PCR) methods. The nucleotide CC sequences can be used in sequencing methods. The present sequence CC represents a reverse transcriptase protein of the invention

XX Sequence 1199 AA;

Alignment Scores:

Pred. No.:	2,016-293	Length:	1199
Score:	3969.50	Matches:	752
Percent Similarity:	76.3%	Conservative:	170
Best Local Similarity:	62.2%	Mismatches:	254
Query Match:	27.1%	Indels:	33
DB:	3	Gaps:	12

US-10-723-552-3 (1-8132) x AAB12994 (1-1199)

QY 2160 GGGAGACGGGTTTCGGACCCCTCCCGAGCCCGAGGTAACCTTTGAAGGTGGAGGGCAA 2219
 Db 1 GlyGlyGlnGlnGlnProProGluProArgIleThrLeuLysValGlyGlyGln 20
 QY 2220 CCAGTTGAGTTCTTGGTTCATACCGGAGGGAACATTCAGTCTACTACAGCCATTAGA 2279
 Db 21 ProValThrPheLeuValAspThrGlyAlaGlnHisSerValLeuThrGlnAsnProGly 40
 QY 2280 AAAATAAAGATATAAAATCTCGGGTGATGGTGGCCACAGGSCAACACAGTATCCATGG 2339
 Db 41 ProLeuSerAspLysSerAlaIleValGlnGlnAlaThrGlyLysArgIleThrArgTrp 60
 QY 2340 ACTACCCGAAGAACAGTTGACTTGGGAGTGGGACGGGTAAACCCACTCGTTCTTGGTCAFA 2399
 Db 61 ThrThrAspArgLysValHisLeuAlaThrGlyLysValIleThrHisSerPheLeuHisVal 80
 QY 2400 CCTGAGTCCGACGACCCCTCTTAGTAGAGACTTATTCACCAAGATGGGACACAAATT 2459
 Db 81 ProAspCysProTyProLeuLeuGlyArgAspLeuLeuThrLysLeuLysAlaGlnIle 100
 QY 2460 TCTTTTGA---CAAGGGAAACAGAAAGTGTCTGCAATAACAAACCTATCATCTGTGTG 2516
 Db 101 HisPheGluGlySerGlyAlaGlnValMetGlyProMetGlyGlnProLeuGlnValLeu 120
 QY 2517 ACCCTCCAAATTAGATGACAAATATCGACTATATCTCCCTAGTAAAGCCTGATCAAAAT 2576
 Db 121 ThrLeuAsnIleGluAspGluHisArgLeuLeuHisGluThrSerLysGluProAspValSer 140
 QY 2577 ATA---CAATTCTGGTTGAAACAGTTTCCCAAGCTGGGACAGAAACCGCAGGATGGGT 2633
 Db 141 LeuGlySerThrThrLeuSerAspPheProGlnAlaIleProAlaGluThrGlyGlyMetGly 160
 QY 2634 TTGGCAAGCAAGTTCCCCCAACAGTTATTCAACTGAAGCCAGTGCACACACAGTCA 2693
 Db 161 LeuAlaValArgGlnAlaProLeuIleIleProLeuLysAlaThrSerThrProValSer 180
 QY 2694 GTACAGACAGTACCCCTTAGTAAAGAGCTCAAGAGGAATTCGGCCGCGATCTCCAAAGA 2753
 Db 181 IleLysGlnTyProMetSerGlnGluAlaArgLeuGlyIleLysProHisIleGlnArg 200
 QY 2754 TTAATCCAAACAGGGCATCTAGTCTCTGTCATCTCCCTGGAATATCTCCCTGCTACCG 2813
 Db 201 LeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeuLeuPro 220
 QY 2814 GTTACAAAGCCCTGGGCAATGACTATCGACCACTACAGGACTTCAGAGAGGTCAATAAA 2873
 Db 221 ValLysLysProGlyThrAsnAspTyArgProValGlnAspLeuArgGlnValAsnLys 240

QY 2874 CCGGTGCAGGATATACACCCAGTCCGAAACCTTATAACCTCTGTGTCTCTCCCA 2933
 Db 241 ArgValGluAspIleHisProThrValProAsnProTyAsnLeuLeuSerGlyLeuPro 260
 QY 2934 CCCCACGAGCTGGTATACAGTATTGGACTTAAAGGATCCCTTCTCTCTCCCTGAGATTA 2993
 Db 261 ProSerHisGlnTrpTyThrValLeuAspLeuLysAspAlaPheCysLeuArgLeu 280
 QY 2994 CACCCCACTAGCAACCACTTTTTCGCTTCGAATGGAGAGATCCAGGTACGGGAAGAACA 3053
 Db 281 HisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGluMetGlyIleSer 300
 QY 3054 GGGCAGCTCCTCGGACCCGACTGCCCAAGGTTCAAGAACTCCCGACCACTCTTTGAC 3113
 Db 301 GlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAsp 320
 QY 3114 GAAGCCCTACACAGAGCTGGCCCACTTCAGATTCAGATTCACACCCCTCAGGTGACCTCTC 3173
 Db 321 GluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuIleLeuLeu 340
 QY 3174 CAGTACGTGGATGACTCTCTCGCGGAGCCACCAACACGAGCTGTCTAGAAAGCACG 3233
 Db 341 GlnTyValAspAspLeuLeuAlaThrSerGluLeuAspCysGlnGlnGlyThr 360
 QY 3234 AAGGCACACTCTCTGGAATTGTCTGACCTAGCTACAGAGCCTCTCTTAAGAGGCCACG 3293
 Db 361 ArgAlaLeuLeuGlnThrLeuGlyAsnLeuGlyTyArgAlaSerAlaLysLysAlaGln 380
 QY 3294 ATTTCGAGGAGAGAGTAACTACTTGGGTACAGTTTTCGGGACGGGACGGGATGGCTG 3353
 Db 381 IleCysGlnLysGlnValLysTyLeuGlyTyLeuLeuLysGluGlyGlnArgTrpLeu 400
 QY 3354 ACGGAGGACGAGCAAGAAATCTAGTCCAGATACCGGCCCAACACACACCAAAATG 3413
 Db 401 ThrGluAlaArgLysGlnThrValMetGlyGlnProProLysThrProArgGlnLeu 420
 QY 3414 AGAGAGTTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGTTTCGACCTTA 3473
 Db 421 ArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGluMet 440
 QY 3474 GCAGCCCACTCTACCCGCTAACCAAGAAAGGGGAATTTCTCTGGGCTCTGAGACAC 3533
 Db 441 AlaAlaProLeuTyProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAspGln 460
 QY 3534 CAGAAGGCATTTGATGCTATCAAAAGCCCTGCTGAGCGCACCTGCTCTGCGCCCTCC 3593
 Db 461 GlnLysAlaTyGlnGlnLysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeuPro 480
 QY 3594 GACGTAACTAAACCCCTTTTACCTTTTATGTGGATGAGCGTAAAGGAGTAGCCCGGGAGTT 3653
 Db 481 AspLeuThrLysProPheGluLeuPheValAspGluLysGlnGlyTyAlaLysGlyVal 500
 QY 3654 TTAACCCAAACCTTAGGACCATGGAGAGACCTGTGCGCTTACTCTGTCAAAAGAGCTCGAT 3713
 Db 501 LeuThrGlnLysLeuGlyProTrpArgProValAlaTyLeuSerLysLysLeuAsp 520
 QY 3714 CCTGTAGCCAGTGGTGGCCCATATGCTCGTGAAGGCTATCGCAGCTGTGGCCATCTGTC 3773
 Db 521 ProValAlaAlaGlyTrpProCysLeuArgMetValAlaAlaIleAlaValLeuThr 540
 QY 3774 AAGCAGCTGACAAATTGACTTTTGGGACAGAAATATAACTGTAATAGCCCCCATGATTG 3833
 Db 541 LysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAlaVal 560
 QY 3834 GAGAACATCGTTGGCAGCCCGACGATGGATGACCAAGCCCGCATGATGCCACTAT 3893
 Db 561 GluAlaLeuValLysGlnProProAspArgTrpLeuSerAsnAlaArgMetThrHisTy 580
 QY 3894 CAAGCCCTGCTCTC---ACAGAGAGGCTCAGTTTCGCTCCACAGCGCTCTCAACCT 3950
 Db 581 GlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProValAlaLeuLeuAsnPro 600
 QY 3951 GCCACTCTCTGCTGAAAGAGACTGTAAGAACAGTGAATCATGATGTCATGATGCAATTG 4010

Db	601	AlaThrLeuLeuPro--LeuProGluGluGlyLeuGlnHisAsnCysLeuAspIleLeu	619
QY	4011	ATTGAGGAGACTGGGTCGCGCAAGACCTTACAGACATACCGCTGACTGGAGAAGTGCTA	4070
Db	620	AlaGluAlaHisGlyThrArgProAspLeuThrAspGlnProLeuProAspAlaAspHis	639
QY	4071	ACCTGGTTCACTGACGGAAGCAGCTATGCTGGTGAAGGTAAAGAGATGCGTGGGGCGCGC	4130
Db	640	ThrTrpTyrThrAspGlySerLeuLeuGlnGluGlyGlnArgLysAlaGlyAlaAla	659
QY	4131	GTGGTGACGGGACCGCCACGATCTGGCCAGCAGACTGCGCGGAAGAACTTCAGCACAA	4190
Db	660	ValThrThrGluThrGluValIleTrpAlaLysAlaLeuProAlaGlyThrSerAlaGln	679
QY	4191	AAGCTCAGCTCATGGCCCTCAGCAAGCTTTCGGCTGCGCGAAGGGAATCCATAAAC	4250
Db	680	ArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysLeuAsn	699
QY	4251	ATTATACGACAGCAGGTATGCTTTGCGACTGCACACGTACATGGGGGCATCTATAAA	4310
Db	700	ValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyGluIleTyrArg	719
QY	4311	CAAAAGGGGTGCTTACTCTCAGCAGGAGGGAATAAAGAACAAAGAGAAATCTTAAGC	4370
Db	720	ArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAspGluIleLeuAla	739
QY	4371	CTATTAGACCGCTACATTTACCAAAAGCTAGCTATTATATACACTCTCTGCACATCAG	4430
Db	740	LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHisGln	759
QY	4431	AAAGCTAAAGATCTCATATACAGAGGAAACAGATGGCTCAGCGGTGGCCAAAGCAGCA	4490
Db	760	LysGlyHisSerAlaGluAlaArgGlyAsnArgMetAlaAspGlnAlaLaaArgLysAla	779
QY	4491	GCCAGGGTGTTAACTTCTGCTTATATAGAAATGCCCAAAGCC-----	4535
Db	780	Ala-----IleThrGluThrProAspThrSerThrLeuLeuIle	792
QY	4536	-----CCNGAACCCAGGACAGTACACCCCTAGNAGACTGCCAAGAG	4577
Db	793	GluAsnSerSerProTyrThrSerGluHisPheHisTyrThrValThrAspIleLysAsp	812
QY	4578	ATAAAAAGATAGACCAAG---TTCTCTGAGACTCCGGAAGGACCTGCTATCTCAGAT	4634
Db	813	LeuThrLysLeuGlyAlaIleTyrAspLysThrLysLysTyrTrpValTyr-----Gln	830
QY	4635	GGGAAGAAATCTTGCCCCCAAAAGAAGGTGTAGAATATGTCCAACAGATACATCGTCTA	4694
Db	831	GlyLysProValMetProAspGlnPheThrPheGluLeuLeuAspPheLeuHisGlnLeu	850
QY	4695	ACCCACCTAGGAATAACACCTGCACAGCTTGGTCAGAACA-----TCCCTTATCAT	4748
Db	851	ThrHisLeuSerPheSerLysMetLeuAlaLeuLeuGluArgSerHisSerProTyrTyr	870
QY	4749	GTTCCTGAGGCTACCAGGAGTGGCTGACTCGTGTGTCAAAATTGTGTGCGCTGCCAGCTG	4808
Db	871	MetLeuAsnArgAspArgThrLeuLysAsnIleThrGluThrCysLysAlaCysAlaGln	890
QY	4809	GTTAATGCTAATCTCCAGATGCCCTCCAGGGAAGAGACTAAAGGGGAAGCCACCAGGC	4868
Db	891	ValAsnAlaSerLysSerAlaValLysGlnGlyThrArgValArgGlyHisArgProGly	910
QY	4869	GCTCCTGGGAAGTGGAATCTTCATGAGGTAAAGCCGGCTAAATACGGAACAATACTCTA	4928
Db	911	ThrHisTrpGluIleAspPheThrGluIleLysProGlyLeuTyrGlyTyrLysTyrLeu	930
QY	4929	TTGGTTTTTGTAGACACTTTCAGGATGGGTAGAGGCTTATCTCTCTAAGAAGAGACT	4988
Db	931	LeuValPheIleAspThrPheSerGlyTyrIleGluAlaPheProThrLysLysGluThr	950
QY	4989	TCAACCTGTGGCTAAAAAATACTCGAAGAAATTTTCCAGATTTGGAATACCTTAAG	5048

Db	951	AlaLysValValThrLysLysLeuLeuGluGluLeuPheProArgPheGlyMetProGln	970
Qy	5049	GTAAATAGGGTCACACAATCGTCCAGCTTTTGTGGCCAGGTAAAGTCAGGACTCGGCCAAG	5108
Db	971	ValLeuGlyThrAspAsnGlyProAlaPheValSerLysValSerGlnThrValAlaAsp	990
Qy	5109	ATATTGGGAGTATTGGAAACTGCATTGTCATATGTCATACAGACCCCAAGCTCAGACAGGTA	5168
Db	991	LeuLeuGlyLeuAspTrpLysLeuHisCysAlaIatyrArgProGlnSerSerGlyGlnVal	1010
Qy	5169	GAGAGATGAATAGAACCATTAAGAGACCCCTTACTAAATTGACCCCGGAGACTGGCGTT	5228
Db	1011	GluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrLeuAlaThrGlySer	1030
Qy	5229	AATGATTGGATAGCTCTCCCTGCCCTTTGTGCTTTTGTAGGGTTAGGAACACCCCTGCAGAC	5288
Db	1031	ArgAspTrpValLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThrProGlyPro	1050
Qy	5289	TTTGGGCTGACCCCTATGAATTACTCTACGGGGGACCCCCCATTTGGTGTAGAAATTGCT	5348
Db	1051	HisGlyLeuThrProTyrGluIleLeuTyrGlyAlaProProProLeuValAsnPhePro	1070
Qy	5349	TCGTGCATAGTCTGACGTGCTGCTTCCACGCCCTTTGTCTCTAGGCTCAAGGCACATT	5408
Db	1071	AspProAspMetThrArgValThrAsnSerProSerLeuGlnAlaHisLeuGlnAlaLeu	1090
Qy	5409	GAGTGGGTGAGACAACGAGCGTGGAGCAACTCCGGGAGGCCCTACTCAGGAGGAGAGAC	5468
Db	1091	TyrLeuValGlnHisGluValTrpArgProLeuAlaAlaTyrGlnGluGlnLeuAsp	1110
Qy	5469	TTGCAG---ATCCCCACATCGTTTCCAAAGTGGGAGATTACGTCTACGTTAGACGCCACCGT	5525
Db	1111	ArgProValValProHisProTyrArgValGlyAspThrValTrpValArgArgHisGln	1130
Qy	5526	GCAGGAAACCTCGAGACTCGGTGGAGGGCCCTTACTCTGTACTTTTGACCAACACCAAG	5585
Db	1131	ThrLysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThrThrProThr	1150
Qy	5586	GCCTGTGAAGTCCGAGGAATCTCCACCTGGATCCATGCATCCCACTTAAACCGGCGCCA	5645
Db	1151	AlaLeuLysValAspGlyIleAlaAlaTrpIleHisAlaAlaHisValLysAlaAlaAsp	1170
Qy	5646	CCT-----CCCATTTCG-----GGGTGGAAGCCGGAAGACTGAAATCCCTT	5690
Db	1171	ProGlyGlyGlyProSerSerArgLeuThrTrpArgValGlnArgSerGlnAsnProLeu	1190
Qy	5691	AAGCTTCGCTCCATCGCTGGTTCCT	5717
Db	1191	LysIleArgLeuThrArgGluAlaPro	1199
RESULT 14			
AA17947			
ID	AA17947 standard; protein; 1224 AA.		
XX	AA17947;		
XX	03-AUG-1999 (first entry)		
DT	MOMLV pol gene product.		
DE	Viral packaging signal; chimeric; type C retrovirus; gag gene; MOMLV;		
XX	murine VL30; retroviral vector; biotechnology; pharmaceutical;		
KW	gene therapy; Moloney murine leukemia virus; pol gene.		
KW	Moloney murine leukemia virus.		
OS	WO925862-A2.		
XX	27-MAY-1999.		
PN	19-NOV-1998; 98WO-US024667.		
XX	19-NOV-1997; 97US-0066148P.		
PF			
XX			
PR			


```
Db 521 ProValAlaAlaGlyTrpProCysLeuArgMetValAlaAlaLeuValLeuThr 540
QY 3774 AAGGACGCTGACAAATGTACTTTGGGACAGAAATATAACTGTAAATAGCCCCCAGCATG 3833
Db 541 LysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAlaVal 560
QY 3834 GAGAAATCGTTCCGAGCCCGACGACCGATGATCAACCAACCCCGCATGACCCACTAT 3893
Db 561 GluAlaLeuValLysGlnProProAspA-gTrpLeuSerAsnAlaArgMetThrHisTyr 580
QY 3894 CAAAGCGCTGCTCTC---ACAGAGAGGGTCACGCTTCGCTCCACGACCGCTCTCAACCCCT 3950
Db 581 GlnAlaLeuLeuLeuAspThrAspA-gValGlnPheGlyProValValAlaLeuAsnPro 600
QY 3951 GCCACTCTTCGCTGAGAGACTGATGAACAGTACTGATCATGATTTGCCATCAACTATTG 4010
Db 601 AlaThrLeuLeuPro---LeuProGluGluGlyLeuGlnHisAsnCysLeuAspIleLeu 619
QY 4011 ATTGAGGAGACTGGGCTCCGACGAGGACCTTACAGACATACCGCTGACTGGAGAGTGCTA 4070
Db 620 AlaGluAlaHisGlyThrArgProAspLeuThrAspGlnProLeuProAspAlaAspHis 639
QY 4071 ACCTGCTTCACTACCGAAGCAGCTATGTGTTGGAAGGTAAAGAGGATGCTGGGGCGCG 4130
Db 640 ThrTrpTyrThrAspGlySerLeuLeuGlnGluGlyGlnArgLysAlaGlyAlaAla 659
QY 4131 GTGGTGGACGGGACCCGACGATCTGGCCGACGAGCTGCGGAGAGAACTTCAGACAA 4190
Db 660 ValThrThrGluThrGluValIleTrpAlaLysAlaLeuProAlaGlyThrSerAlaGln 679
QY 4191 AAGCTGAGCTCATGGCCCTCAGCAGCTTTCGGCTGCGGCTGCCGACGGAATCCATAAC 4250
Db 680 ArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysLysLeuAsn 699
QY 4251 ATTTATACGACAGCAGGTATGCTTTGCGACTGCACATCATCGGGCCATCTATAAA 4310
Db 700 ValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyGluIleTyrArg 719
QY 4311 CAAAGGGGTGTCTACTCAGCAGGGGAGGAAATAAAGAACAAAGAGGAAATCTTAAGC 4370
Db 720 ArgGlyLeuLeuLeuThrSerGluGlyLysGluIleLysAsnLysAspGluIleLeuAla 739
QY 4371 CTATTAGAGCCGTACATTACCAAAAGCTAGCTATTATACATCTCTCGACATCAG 4430
Db 740 LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHisGln 759
QY 4431 AAAGCTAAAGATCTCATATCCAGGAAACACGATGGCTGACCGGTGCGCAAGCAGGCA 4490
Db 760 LysGlyHisSerAlaGluAlaArgGlyAsnArgMetAlaAspGlnAlaAlaArgLysAla 779
QY 4491 GCCAGGGGTGTTAACTTCTGCTTAATAAGAAATGCCAAAGCC----- 4535
Db 780 Ala-----IleThrGluThrProAspThrSerThrLeuLeuIle 792
QY 4536 -----CCAGAACCCAGCAGCAGTACACCTCAGAACGATGCGCAAGAG 4577
Db 793 GluAsnSerSerProTyrThrSerGluHisPheHisTyrValThrAspIleLysAsp 812
QY 4578 ATAAAAAAGATAGACCAAG---TTCTCTGACATCCCGAGGGACCTGTCTATACCTCAGAT 4634
Db 813 LeuThrLysLeuGlyAlaIleTyrAspLysThrLysLysTyrTrpValTyr-----Gln 830
QY 4635 GGAAGGAAATCTGCGCCCAAGAGGGTTAGAAATATGTCACAGATACATCTCTA 4694
Db 831 GlyLysProValMetProAspGlnPheThrPheGluLeuLeuAspPheLeuHisGlnLeu 850
QY 4695 ACCCATAGGAACCTAAACACTGCGACGAGTTGGTCAGACA-----TCCCTTATCAT 4748
Db 851 ThrHisLeuSerPheSerLysMetLysAlaLeuLeuGluArgSerHisSerProTyrTyr 870
QY 4749 GTTCTCAGGCTACAGGAGTGCTGACTCGGTGGTCAACATTTGTGCGCTCCAGCTG 4808
Db 870
```

```
871 MetLeuAsnArgAspArgThrLeuLysAsnIleThrGluThrCysLysAlaCysAlaGln 890
QY 4809 GTTAATGCTAATTCCTTCCAGAATGCTCCAGGAGAGACTAAGGGAAGCCACCCAGGC 4868
Db 891 ValAsnAlaSerLysSerAlaValLysGlnGlyThrArgValArgGlyHisArgProGly 910
QY 4869 GCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAACAAATACCTA 4928
Db 911 ThrHisTrpGluIleAspPheThrGluIleLysProGlyLeuTyrGlyTyrLysTyrLeu 930
QY 4929 TTGGTTTTGTAGACACCTTTTTCAGGATGGGTAGAGCTTATCTCTACTAAGAAAGACT 4988
Db 931 LeuValPheIleAspThrPheSerGlyTyrIleGluAlaPheProThrLysLysGluThr 950
QY 4989 TCAACGCTGGTGTAAAAAATACTCGGAAGAAATTTTCCAGATTTCCGAATACCTAAG 5048
Db 951 AlaLysValValThrLysLysLeuLeuGluIlePheProArgPheGlyMetProGln 970
QY 5049 GTAATAGGCTCAGACAATGCTCAGCTTTTGTGCCCAGCTAAGTCAGGACTGGCCAAAG 5108
Db 971 ValLeuGlyThrAspAsnGlyProAlaPheValSerLysValSerGlnThrValAlaAsp 990
QY 5109 ATATTGGGATTCATTGGAACTGTCATTGTGCATACAGACCCCAAGCTCAGGACAGGTA 5168
Db 991 LeuLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnVal 1010
QY 5169 GAGAGATCAATAGAACCATTAAGAGACCTTACTTAATTTGACCCGCGAGACTGCGGT 5228
Db 1011 GluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrLeuAlaThrGlySer 1030
QY 5229 AATGATGATGATGCTCTCCTGCTCTTGTGCTTTTGTAGGTTAGGAACACCCCTGACAG 5288
Db 1031 ArgAspTrpValLeuLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThrProGlyPro 1050
QY 5289 TTTGGCTGACCCCTTATGAATTTACTACGGGGGACCCCCCTCATTTGTAGAAATTTGCT 5348
Db 1051 HisGlyLeuThrProTyrGluIleLeuTyrGlyAlaProProProLeuValAsnPhePro 1070
QY 5349 TCTGTACATAGTGTGCTGCTGCTGCTTTTCCAGCCTTTGCTCTAGGCTCAAGGCATT 5408
Db 1071 AspProAspMetThrArgValThrAsnSerProSerLeuGlnAlaHisLeuGlnAlaLeu 1090
QY 5409 GAGTGGCTCAGACAACAGCGCTGGAGGCAACTCCGGAGGCTCTCTCAGGAGGAGGAGAC 5468
Db 1091 TyrLeuValGlnHisGluValTrpArgProLeuAlaAlaTyrGlnGlnLeuAsp 1110
QY 5469 TTGCAG---ATCCACATCGTTTCCAAAGTGGGAGATTTCAGTCTACTAGACCCACCGT 5525
Db 1111 ArgProValValProHisProTyrArgValGlyAspThrValTrpValArgArgHisGln 1130
QY 5526 GCAGGAACTCGAGACTCGGTGGAGGGCCCTTATCTCGTACTTTTGTGACCAACCAACG 5585
Db 1131 ThrLysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThrThrProThr 1150
QY 5586 GCTGTGAAAGTCGAAGGAATCTCCACTCGATCCATGCTATCCACGTTAAACCGCGCCA 5645
Db 1151 AlaLeuLysValAspGlyIleAlaIleAlaIleHisAlaHisValLysAlaAlaAsp 1170
QY 5646 CTTCCGATTCG----- 5657
Db 1171 ProGlyGlyProMetAlaArgSerThrLeuSerLysProLeuLysAsnLysValAsn 1190
QY 5658 -----GGGTGGAAGCCGAAAG 5675
Db 1191 ProArgGlyProLeuLeuProLeuIleLeuSerSerArgLeuThrTrpArgValGlnArg 1210
QY 5676 ACTGAAATCCCTTAAAGCTTCGCTCCATCCGCTGCTTCCCT 5717
Db 1211 SerGlnAsnProLeuLysIleArgLeuThrArgGluAlaPro 1224
Db 1224
```

RESULT 15
ADS73450
ID ADS73450 standard; protein; 744 AA.

Qy	1263	AACCATCCCTTCTCGAGGATCCCAAGCCTCACGGGTTGGTGGAGTCCCTTATG	1322
Db	447	AsnHisProProPheSerGluAspProGlnArgLeuThrGlyLeuValGluSerLeuMet	466
Qy	1323	TTCTCTCACAGCCTACTTGGGATGATTGTCAACAGCTGTGCAGACACTCTTCACAACC	1382
Db	467	PheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheThrThr	486
Qy	1383	GAGGAGCGAGAGAAATCTATTAGAGCTAGAAAAATGTTCTTGGGGCCGACGGGCGA	1442
Db	487	GluGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAlaAspGlyArg	506
Qy	1443	CCACAGCGGTGCNAATGAGATTGACATGGGATTCCTTAACCTGCCCGGTTGGGAC	1502
Db	507	ProThrArgLeuGlnAsnGluIleAspMetGlyPheProLeuThrArgProGlyTrpAsp	526
Qy	1503	TACAACACCGCTGAAGTAGGAGAGCTTCAAAATCTATCGCAGAGCTCTGGTGGCGGT	1562
Db	527	TyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArgGlnAlaLeuValAlaGly	546
Qy	1563	CTCGGGGCGCTCAAGACGGCCCACTAATTGGCTAAGCTAAGAGAAAGTGATCGAGGA	1622
Db	547	LeuArgGlyAlaSerArgProThrAsnLeuAlaLysValArgGluValMetGlnGly	566
Qy	1623	CCGAATGAACCCCTCTGTTTCTTGAGAGGCTCTTGAAGCCTTCAGGCGGTACACC	1682
Db	567	ProAsnGluProProSerValPheLeuGluArgLeuLeuGluAlaPheArgArgTyrThr	586
Qy	1683	CTTTTTCATCCACTCAGAGGCGCCAAAAGCCTCAGTGGCTTTGGCTTTATAGACAG	1742
Db	587	ProPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheIleGlyGln	606
Qy	1743	TCAGCCTTGGATATTGAAAGAGCTTCAGAGACTGGAAGGTTACAGGAGGCTGAGTTA	1802
Db	607	SerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGlyLeuGlnGluAlaGluLeu	626
Qy	1803	CGTGATCTAGTGAAGGCGCAGAGAAAGTATATTACAAAGGAGACAGAGAAGAAAGG	1862
Db	627	ArgAspLeuValLysGluAlaGluLysValTyrTyrLysArgGluThrGluGluArg	646
Qy	1863	GAACAAAGAAAAGAGAGAGAAAGAGAGGAGGAGAAAGACGTATATAACGGCAAGAG	1922
Db	647	GluGlnArgLysGluArgGluArgGluArgGluArgGluArgAsnLysArgGlnGlu	666
Qy	1923	AGAATTGACTAAGATCTTGGCTGCTGAGTGGTTGAAGGGAAGCAATACGGAAGAGAG	1982
Db	667	LysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLysSerAsnThrGluArgGlu	686
Qy	1983	AGAGATTTTAGGAAAATTAGGTCAGGCCCTAGACAGTCAGGGAACCTGGGCAATAGGACC	2042
Db	687	ArgAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGlyAsnArgThr	706
Qy	2043	CCACTCGAAGGACCAATGTGCATATTGTAAGAAAGAGACACTGGCAAGGAACCTGC	2102
Db	707	ProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGlyHisTrpAlaArgAsnCys	726
Qy	2103	CCCAAGAGGGAAACAAAGACCAAGGATCCTAGCTCTAAGAAAGATAAAGAT	2156
Db	727	ProLysLysGlyAsnLysGlyProArgIleLeuAlaLeuGluAspLysAsp	744

Search completed: February 14, 2006, 15:50:14
Job time : 1473.8 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 15:55:36 ; Search time 42.0295 Seconds
(without alignments)
3199.272 Million cell updates/sec

Title: US-10-723-552-3
Perfect score: 14636
Sequence: 1 GCGTGTGTACGACTGTGG.....CTGTTTGCATCAAAAAA 8132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DBV=xlp
-O=/abes/ABSSWEB spool/US10723552/runat_14022006_125146_13079/app_query.fasta.1
-DB=Issued Patents.AA -QFMT=fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs802p -USER=US10723552 @CN 1 1 161 @runat_14022006_125146_13079
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6099.5	41.7	1194	2	US-09-171-553B-5
2	5338	36.5	1737	2	US-09-309-572-13
3	5338	36.5	1737	2	US-09-718-096-13
4	4329.5	29.6	1203	2	US-09-075-272-4
5	3600	24.6	1079	1	US-08-929-967-8
6	3425	23.4	638	2	US-09-376-781-6
7	2866.5	19.6	660	2	US-09-111-085-2
8	2866.5	19.6	660	2	US-09-376-781-5
9	2640	18.0	524	2	US-09-171-553B-4
10	2500.5	17.1	665	1	US-08-929-967-7
11	2422	16.5	657	2	US-09-111-085-4
12	2422	16.5	657	2	US-09-376-781-4

13	2395	16.4	657	2	US-09-171-553B-10
14	2353	16.1	656	2	US-09-376-781-3
15	2353	16.1	656	2	US-09-171-553B-6
16	2297	15.7	1312	2	US-09-554-572-26
17	1796.5	12.3	622	2	US-09-075-272-2
18	1793	11.8	522	2	US-09-075-272-3
19	1505.5	10.3	667	2	US-09-315-127-5
20	1505.5	10.3	667	2	US-09-315-127-6
21	1505.5	10.3	667	2	US-09-070-630-13
22	1501.5	10.3	379	2	US-09-603-185-6
23	1499.5	10.2	768	2	US-08-979-847B-89
24	1477.5	10.1	673	2	US-09-075-272-5
25	1390.5	9.5	683	2	US-08-979-847B-198
26	1390.5	9.5	683	2	US-08-979-847B-200
27	1387.5	9.5	683	2	US-08-979-847B-208
28	1387.5	9.5	683	2	US-08-979-847B-210
29	1380.5	9.4	538	2	US-09-309-572-12
30	1380.5	9.4	538	2	US-09-718-096-12
31	1379.5	9.4	538	2	US-09-370-368-9
32	1331.5	9.1	654	2	US-08-979-847B-91
33	1320	9.0	665	2	US-09-309-572-14
34	1320	9.0	665	2	US-09-718-096-14
35	1318.5	9.0	632	2	US-09-315-127-2
36	1318.5	9.0	632	2	US-09-315-127-3
37	1303.5	8.9	768	2	US-09-120-653D-5
38	1277	8.7	645	2	US-09-315-127-8
39	1277	8.7	645	2	US-09-315-127-9
40	1262.5	8.6	654	2	US-09-315-127-11
41	1262.5	8.6	654	2	US-09-315-127-12
42	692.5	4.7	469	1	US-08-484-126-1
43	692.5	4.7	469	2	US-09-374-909-1
44	660.5	4.5	445	1	US-08-447-925-6
45	639.5	4.4	453	1	US-08-484-126-3

ALLEGMENTS

RESULT 1
US-09-171-553B-5
Sequence 5, Application US/09171553B
Patent No. 6756227
GENERAL INFORMATION:
APPLICANT: GALBRAITH, DANIEL N.
APPLICANT: HAWORTH, CHRISTINE
APPLICANT: LEBES, GILLIAN M.
APPLICANT: SMITH, KENNETH T.
TITLE OF INVENTION: PORCINE RETROVIRUS
FILE REFERENCE: CFV-5.01
CURRENT APPLICATION NUMBER: US/09/171,553B
CURRENT FILING DATE: 1999-02-08
PRIOR APPLICATION NUMBER: PCT/GB97/01087
PRIOR FILING DATE: 1997-04-18
PRIOR APPLICATION NUMBER: GB 9702668.6
PRIOR FILING DATE: 1997-02-10
PRIOR APPLICATION NUMBER: GB 9608164.1
PRIOR FILING DATE: 1996-04-19
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1194
TYPE: PRT
ORGANISM: Porcine retrovirus
US-09-171-553B-5

Alignment Scores:
Pred. No.: 0
Score: 6099.50
Percent Similarity: 98.5%
Best Local Similarity: 96.9%
Query Match: 41.7%
DB: 2
Length: 1194
Matches: 1154
Conservative: 19
Mismatch: 17
Indels: 1
Gaps: 1

US-10-723-552-3 (1-8132) x US-09-171-553B-5 (1-1194)

QY 2160 GGGAGCGGGTTCGGACCCCTCCCGAGCCAGGCTAACTTTGAAGTGGAGGGCAA 2219
DB |||||
1 GlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGln 20
QY 2220 CCAGTTGAGTTCCTGGTTGATACCGGAGCGAAACATTACAGTGTCTACTACAGCCATTAGGA 2279
DB |||||
21 ProValGluPheLeuValAspThrGlyAlaGluHisSerValLeuLeuGlnProLeuGly 40
QY 2280 AAATAAAGATAAAAATCCTGGGTGATGGGTGCCACAGGGCAACAAAGTATCCATGG 2339
DB |||||
41 LysLeuLysGluLysSerTrpValMetGlyAlaThrGlyGlnArgGlnTrpProTrp 60
QY 2340 ACTACCCGAGAACAGTGTACTTGGGAGTGGGCGGTACCCACTCGTTCTCTGGTCATA 2399
DB |||||
61 ThrThrArgArgThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValIle 80
QY 2400 CTTGAGTGGCCAGCACCCCTCTTAGGTAGAGACTTATTGACCAGATGGGAGACAAAT 2459
DB |||||
81 ProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnIle 100
QY 2460 TCTTTGAAACAGGGAACCCAGAGTGTCTGCAATAACAAACCTATCACTGTGTGACC 2519
DB |||||
101 SerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleThrValLeuThr 120
QY 2520 CTCCAATTAGATGACGAATATGACTATCTCTCCCTAGTAAAGCCTGATCAAAATATA 2579
DB |||||
121 LeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnValLysProAspGlnAspIle 140
QY 2580 CAATTCTGGTTGGAACAGTTTCCCAAGCTGGGCAGAAACCCGAGGATGGGTTGGCA 2639
DB |||||
141 GlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuAla 160
QY 2640 AAGCAAGTTCCTCCACAAAGTTATTCAACTGAAGCCAGTCCACACCAAGTGTCACTCAGA 2699
DB |||||
161 LysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValArg 180
QY 2700 CAGTACCCCTTGAGTAAGAGCTCAAGAGGAATTCGGCCGCTATGCCAAGATTAAATC 2759
DB |||||
181 GlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuIle 200
QY 2760 CAACAGGCACTCTAGTTCCTGTCCAATCTCCCTGGAATACTCCCTGCTACCGGTAGA 2819
DB |||||
201 GlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValArg 220
QY 2820 AAGCTGGGACTAATGACTATGACACCAAGTACAGACTTGAGAGAGTCAATAAACGGGTG 2879
DB |||||
221 LysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgVal 240
QY 2880 CAGGATATACCCCAACAGTCCCGAACCTTATACCTTGTGTGCTCTCCACCCCAA 2939
DB |||||
241 GlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSerAlaLeuProProGlu 260
QY 2940 CGGAGCTGGTATACAGTATTGGACTTAAAGATGCCCTTCTTCTGCTGAGATTACACCC 2999
DB |||||
261 ArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisPro 280
QY 3000 ACTAGCCAACCACTTTTGGCTTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGSCAG 3059
DB |||||
281 ThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyGln 300
QY 3060 CTACCTGGAGCCGACTGCCCAAGGTTCAAGAACTCCCGGACCATCTTTTGAGGAAGCC 3119
DB |||||
301 LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluAla 320
QY 3120 CTACAGAGACCTGGCCAACTTCAGGATCCAAACACCTCAGGTAGCCCTCTCCAGTAC 3179
DB |||||
321 LeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnTyr 340
QY 3180 GTGGATGACCTGCTTCTGGCGGGAGCCCAACCAAGGACTGCTTAGAAGGCACCAAGGCA 3239
DB |||||
341 ValAspAspLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysAla 360

QY 3240 CTACTGCTGGAATTGTCTGACCTAGCTAGCTACAGAGCCTCTGCTAAGAGGCCAGATTTCG 3299
DB |||||
361 LeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleCys 380
QY 3300 AGGAGAGAGGTAAACATACTTTGGGGTACAGTTTCGGGACCGGCGAGCATGGCTGACGGAG 3359
DB |||||
381 ArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrGlu 400
QY 3360 GCACGGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACACGCAACAAATGAGAGAG 3419
DB |||||
401 AlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGlnValArgGlu 420
QY 3420 TTTTTCGGGACAGCTGATTTTCAGACTGTGTGATCCCGGGTTTCGACCTTAGCAGGCC 3479
DB |||||
421 PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAla 440
QY 3480 CCACTCTACCCCGCTAAACCAAGAAAAAGGGGAAATTCCTCTGGGCTCTCTGAGCACCAAG 3539
DB |||||
441 ProLeuTyrProLeuThrLysGluLysGlyGlyPheSerTrpAlaProGluHisGlnLys 460
QY 3540 GCATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGAGGTA 3599
DB |||||
461 AlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspVal 480
QY 3600 ACTAAACCCCTTTACCCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGGGAGTTTAAAC 3659
DB |||||
481 ThrLysProPheThrLeuTyrValAspGluLysGlyValAlaArgGlyValLeuThr 500
QY 3660 CAAACCTTAGGACCATGGAAGACCTGTGCGCTACTCTGTCAAAGAGCTCGATCTCTGTA 3719
DB |||||
501 GlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProVal 520
QY 3720 GCAGTGGTTGGCCCATATGCCGTAAGGCTATGCGAGCTGTGCGCCATCTGCTCAAGGAC 3779
DB |||||
521 AlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLeuValLysAsp 540
QY 3780 GCTGACAAATGTACTTTGGACAGATAAATGTAATAGCCCCCATGCTGATGGAGAAC 3839
DB |||||
541 AlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluAsn 560
QY 3840 ATCGTTCCGCGAGCCCGCAGACCGATGATGACCAACGCCCGCATGACCCCATATCAAGC 3899
DB |||||
561 IleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnSer 580
QY 3900 CTGCTTCTACAGAGAGGTTCAGTTTCGCTCCACGAGCCGCTCTCAACCTCGCATCTTT 3959
DB |||||
581 LeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrLeu 600
QY 3960 CTGCTGAAAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTATTGATTGAGGAG 4019
DB |||||
601 LeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluGlu 620
QY 4020 ACTGGGGTCGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGTAACTGCTTTC 4079
DB |||||
621 ThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpPhe 640
QY 4080 ACTGACGGAAGCAGCTATGTGTGGAGAGGTAAAGAGATGCTGGGGGGGGTGGGAC 4139
DB |||||
641 ThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGlyAlaAlaValValAsp 660
QY 4140 GGCACCCGACGACTCTGGGCGCAGCAGCTCGCGGAAGGAACTTTCAGCACAAAAGGCTGAG 4199
DB |||||
661 GlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaGlu 680
QY 4200 CTCATGGCCCTCAACGCAAGCTTTTGGCTGGCCGGAAGGGAAATCCATAAACATTATACG 4259
DB |||||
681 LeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrThr 700
QY 4260 GACAGAGGTATGCCCTTTGGACTGACACAGCTACATGGGGCCATCTATAAACAAAGGGG 4319
DB |||||
701 AspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgGly 720
QY 4320 TTGCTTACCTCAGCAGGGGGGAAATAAAGAAACAAAGAGGAAATTCCTAAGCCTATTAGAA 4379

Db 721 LeuLeuThrSerAlaGlyArgGluLeuLysGluLeuLysGluLeuSerLeuLeuGlu 740
QY 4380 GCGCTACATTTACCAAAAGCGTAGCTATTATACACTGTCTCGACATCAGAAAGCTAAA 4439
Db 741 AlaLeuHisLeuProLysArgLeuAlaIleHisCysProGlyHisGlnLysAlaLys 760
QY 4440 GATCTCATATCCAGAGAAACCGATGGCTGACGGGTTCCCAAGCAGGCGCCCGGT 4499
Db 761 AspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaGlnAla 780
QY 4500 GTTAACTCTGCTCTAATAGAAATGCCCAAGCCCAAGACCCAGACCCAGACAGTACACC 4559
Db 781 ValAsnLeuLeuProIleLeuThrProLysAlaProGluProArgArgGlnLysThr 800
QY 4560 CTAGAAGACTGGCAAGAGATAAAAGATAGACAGTCTCTCAGACTCCCGGAAGGAGCC 4619
Db 801 LeuGluAspTrpGlnGluLeuLysLysLysLysLysLysLysLysLysLysLysLys 820
QY 4620 TGCCTATACCTCAGATGGGAAGAAATCTCTCCCAAGAAAGGGTTAGATATGTCCTAA 4679
Db 821 CysTyThrSerTyrgLysGluLeuLeuProHisLysGluGlyLeuGluTyValGln 840
QY 4680 CAGATACATGCTTAACCCACCTAGGACTAGAACTAAACACTGCAGCAGTGTGCAGAACATCC 4739
Db 841 GlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrSer 860
QY 4740 CTTTATCATGTTCTGAGGCTACCAAGAGTGGCTGACTCGGTGTCAAACTGTGTGCC 4799
Db 861 ProTyHisValLeuArgLeuProGlyValAlaAspSerValLysHisCysValPro 880
QY 4800 TGCACGCTGTTAATCTAATCTCTCAGAAATCCCTCAGGGAAGAGACTAAGGGGAAGC 4859
Db 881 CysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlySer 900
QY 4860 CACCCAGGCTCCTACCTGGAGTGGACTTCACTGAGGTAAAGCCGCTAAATACGAAAC 4919
Db 901 HisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLysTyrgLys 920
QY 4920 AAATACCTATTGTTTGTAGACACTTTCAGGATGGTAGAGCTTACTCTACTAAG 4979
Db 921 LysTyLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyProThrLys 940
QY 4980 AAGAGACTCAACCTGTGGTGTAAATAAATACTCGAAGAAATTTTCCAGATTGGA 5039
Db 941 LysGluThrSerThrValAlaLysLysLysLysLysLysLysLysLysLysLysLys 960
QY 5040 ATACCTAAGGTAAAGGTACAGACATGTCTCAGCTTTTGTTCCTCCAGTAAAGTACGGA 5099
Db 961 IleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnGly 980
QY 5100 CTGGCCCAAGATATTGGGATTTGATTCGAAACTGCTATTGTGCATACAGACCCCAAGCTCA 5159
Db 981 LeuAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1000
QY 5160 GCACAGGTAGAGAGATGAATAGAACATTAAGAGACCTTACTAAATTTGACCGGGAG 5219
Db 1001 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLysLeuThrGlu 1020
QY 5220 ACTGGCGTTAATGATTGATAGCTCTCTCCGCTTTGTGCTTTTGTAGGGTTAGGAACACC 5279
Db 1021 ThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnThr 1040
QY 5280 CTTGGACAGTTTGGGCTGACCCCTATGAATTAATCTACGGGGGACCCCTCCCTTGGTA 5339
Db 1041 ProGlyGlnPheGlyLeuThrProTyrgLysLeuLeuLeuTyrgLysProProLeuAla 1060
QY 5340 GAAATGCTCTCTCATAGTGTGACGTGCTGCTTCTCCAGCTTTGTCTCTAGGCTC 5399
Db 1061 GluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgLeu 1080
QY 5400 AAGGCACTTGTGGGTAGACAACGAGCGTGGAGGCAACTCCGGAGGCGCTTACTCAGGA 5459

Db 1081 LysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTySer--- 1099
QY 5460 GGAGGAGACTGTCAGATCCACATCGTTTCCAGTGGGAGATTACAGTCTACGTTAGACGC 5519
Db 1100 GlyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyValArgArg 1119
QY 5520 CACCGTGCAGGAAACCTCGAGACTCGGTGGAGGCGCTTATCTCGTACTTTTGCACACA 5579
Db 1120 HisArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrgLysLeuValLeuThrThr 1139
QY 5580 CCAACCGCTGTGAAGTCGNAAGTAATCTCCACTGCATCCATCCACCTGATTAACCG 5639
Db 1140 ProThrAlaValLysValGluGlyIleProThrTrpIleHisAlaSerHisValLysPro 1159
QY 5640 GCCCACCCTCCCGATTCGGGTGAAAGCCGAAAGACTGAAAATCCCTTAAAGCTTCGC 5699
Db 1160 AlaProProAspSerGlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArg 1179
QY 5700 CTCCATCGGTGTTCTTCTTACTCTGTCAATAAC 5732
Db 1180 LeuHisArgLeuValProTyrgLysAsnAsn 1190
RESULT 2
US-09-309-572-13
; Sequence 13, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag-pol protein
US-09-309-572-13
Alignment Scores:
Pred. No.: 0 Length: 1737
Score: 5338.00 Matches: 1035
Percent Similarity: 71.6% Conservative: 239
Best Local Similarity: 58.2% Mismatches: 395
Query Match: 36.5% Indels: 110
DB: 2 Gaps: 26
US-10-723-552-3 (1-8132) x US-09-309-572-13 (1-1737)
QY 585 ATGGGACAGACGGTGACGACCCCTCTTACTTTGACTCTCGACCATTTGAGTAA 644
Db 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
QY 645 TCAGGGCTCATTAATTTGTCAGTTAGAGGGACCTTGGCAGACTTTCTCTGTC 704
Db 21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgTrpValThrPheCysSer 40
QY 705 TCTGAATGCGCCACATTCGATCTGATGCGCATCAGAGGGACCTTTAATCTGAGATT 764
Db 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
QY 765 ATCTGGCTGTTAAAGCAGTATTATTTTCAGACTGACCGGCTCTCATCCGATCAGGAG 824
Db 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
QY 825 CCTATATCTTACCTGAGCAAGATTGGCAGAGGATCTCCGCCATGGGTTAAACATGG 884
Db 81 ProTyIleValThrTrpGluAlaLeuAlaPheAspProProTrpValLysProPhe 100

QY 885 CTGAATAAGCCAAAGAGCCAGGTCCTCCCGAATTTCTGGCTCTTGGAGAGAAACAAACAC 944
Db 101 ValHis----ProLysProProProProLeuProProSerAlaProSerLeuProLeuGlu 119
QY 945 TCGGCTGAATAAGTCACAGCCCTCTCTCATATCTACCCCGAGATGAGGACCA---CCG 1001
Db 120 ProProArgSerThrProProArgSerSerLeuTyrProAlaLeuThrProSerLeuGly 139
QY 1002 GCTTGGCCGGAACCCCA-----TCT 1022
Db 140 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuIleAspLeuLeuThr 159
QY 1023 GTTCCCCACCCCTTATCTGCGACAGGTCGCGGAGGAGCCCTTGTGCCCTCTCT--- 1079
Db 160 GluAspProProProTyr-----ArgAspProArgProProProSer 173
QY 1080 -----GGAGCTCCGCGGTGAGGGACCT----- 1103
Db 174 AspArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyGluAlaProAspProSerPro 193
QY 1104 ---GTCAGGGACTCGAGCCGAGGGGCGCCACCCCGGAGCGGACAGAGATCGCG 1160
Db 194 MetAlaSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrThrSerGln 213
QY 1161 ACATTACCGCTGCGCAGTACGCGCCCTCCACACCGGGGGCCCAATTGACAGCCCTCCAG 1220
Db 214 AlaPheProLeuArgAlaGlyGly-----AsnGlyGln-----LeuGln 226
QY 1221 TATTGGCCCTTTCTCTCGAGATCTCTATAATTGGAATACTAACCATCCCCCTTTCTCG 1280
Db 227 TyrTrpProPheSerSerSerAspLeuTyrAsnTrpLysAsnAsnAsnProSerPheSer 246
QY 1281 GAGATCCCAACCGCTCAGCGGGTGTGTGAGTTCCTTATGTTCTCTACACGCTACT 1340
Db 247 GluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuIleThrHisGlnProThr 266
QY 1341 TGGGATCATGTCAACAGCTGCTGCAGACACTTTCACACCGGAGCGGAGAGAGATT 1400
Db 267 TrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluLysGlnAlaGval 286
QY 1401 CTATTAGAGCTAGAAAAATTTCTTGGGGCGAGCGGCGACCCAGCGGTTCGAAAAAT 1460
Db 287 LeuLeuGluAlaArgLysAlaValArgGlyAspAspGlyArgProThrGlnLeuProAsn 306
QY 1461 GAGATTGACATGGGATTTCCCTTAATCTCGCCCGCGTTGGGACTACACAGCGCTGAAGT 1520
Db 307 GluValAspAlaAlaPheProLeuGluArgProAspTrpAspTyrThrThrGlnAlaGly 326
QY 1521 AGGAGAGCTTGAATCTATCCGAGCTCTGTGCGGGTCTCGGGCGGCTCAAGA 1580
Db 327 ArgAsnHisLeuValHisTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 346
QY 1581 CGGCCCACTAATTTGGCTAAGTAAGAGAGTGTATGAGGAGCGAGTGAACCCCTCT 1640
Db 347 SerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer 366
QY 1641 GTTTTCTTGAGAGCTCTTGGAGCCCTCAGCGGTACACCCCTTTGTATCCACCTCA 1700
Db 367 AlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp 386
QY 1701 GAGGCCCAAAAGCCTCAGTGGCTTTGGCCTTTATAGGACAGTCAGCTTGGATATAGA 1760
Db 387 ProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly 406
QY 1761 AAGAAGCTTCAGACATGGAAGGTTACAGGAGCTGAGTTAGCTGATCTAGTGAAGGAG 1820
Db 407 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 426
QY 1821 GCAGAGAACTATATTACAAAAAGGAGGACAGAGAAGAAAGGAAACAAAGAAAGAGAGA 1880
Db 427 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgGlnIleArgArg 446

QY 1881 GAAAGAGAGAAAGGAGGAAGA-----CGT 1907
Db 447 GluThrGluGluLysGluGluArgArgThrGluAspGluGlnLysGluLysGluArg 466
QY 1908 AATAAAGCGAAGAGAAGAAATTTGACTAAGATCTTGGCTGCACTGTTGAAGGAAAGC 1967
Db 467 AspArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys 486
QY 1968 AATACGGAAGAGAGAGATTTTAGAAAAATTAGTTCAGGCGCTAGACAGTCAGGGAAC 2027
Db 487 Gln-----AspArgGlnGlyGlu 493
QY 2028 CTGGCAATAGGACCCCACTCGCAAGAGACCAATGTGCATATTGTAAGAAAGAGACAC 2087
Db 494 ---ArgArgArgSerGlnLeuAspArgAspGlnCysAlaTyrCysLysGluLysGlyHis 512
QY 2088 TGGCAAGGAACCTGCCCCAAGAG-----GGAACAAGAGCAAGG----- 2129
Db 513 TrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArgProGlnThrSer 532
QY 2130 ATCTTAGCTCTAGAAGAAGATAAAGATTAGGGGAGACGGGTTCCGACCCCTCCCGAG 2189
Db 533 LeuLeuThrLeuAspAsp-----GlyGlyGlnGlyGlnGluProProProGlu 548
QY 2190 CCAGGCTAATTTGAAGGTGGAGGGCAACCAAGTTGAGTTCCTGTTGATACCGAGCG 2249
Db 549 ProArgIleThrLeuLysValGlyGlyGlnProValThrPheLeuValAspThrGlyAla 568
QY 2250 AACATTAGTGTACTACAGCCATTAGGAAACTAAAGATAAAAAATCTCGGTGATG 2309
Db 569 GlnHisSerValLeuThrGlnAsnProGlyProLeuSerAspLysSerAlaTrpValGln 588
QY 2310 GGTGCCACAGGCAACCAACGATATCCATCGACTACCCGAGACACTTGCTTGGAGTG 2369
Db 589 GlyAlaThrGlyGlyLysArgTyrArgTrpThrThrAspArgLysValHisLeuAlaThr 608
QY 2370 GAGCGGTAAACCACTCTTCTGTCATACCTGAGTGCACGACCCCTCTTAGGTAGA 2429
Db 609 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 628
QY 2430 GACTTATTGACCAAGATGGGAGCACAAATTTCTTTTGA---CAAGGGAAACCAAGATG 2486
Db 629 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValMet 648
QY 2487 TCTGCAATAACAAACCTATCTGTTGACCCCTCAATTAGATCAGCAATATCGACTA 2546
Db 649 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu 668
QY 2547 TACTCTCCCTAGTAAAGCTGATCAAAATATA---CAATTCTGTTGGAACAGTTTCCC 2603
Db 669 HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 688
QY 2604 CAAGCTTGCGCAGAAAACCGAGGATGGGTTTGGCAAGCAAGTCTCCCCACAAGTTATT 2663
Db 689 GlnAlaTrpAlaGluThrGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 708
QY 2664 CAACTGAAGCCAGTCCACACCGAGTCAGTCAGACAGTACCCCTTGATTAAGAGCT 2723
Db 709 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrPrometSerGlnGluAla 728
QY 2724 CAAGAAGGAATTCGGCGCATGTCAAAAGATTAAATCAACAGGGCATCTAGTTCTCTGTC 2783
Db 729 ArgLeuGlyLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 748
QY 2784 CAATCTCCCTGGAAATCTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGCTATCGA 2843
Db 749 GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 768
QY 2844 CAGTACAGCACTTGAGAGAGTCAATAAACCGGTGACGATATACACCAACAGTCCCG 2903
Db 769 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 788
QY 2904 AACCCCTTAAACCTCTTGTGTCTCTCCCAACCCCAACGGAGCTGGTATACAGTATTGGAC 2963

Db 789 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp 808
QY 2964 TTAAGAGTACCTCTCTGCTGAGATTACACCCACTAGCCCAACACACTTTTGGCTTC 3023
Db 809 LeuLysAspAlaPhePheCysLeuAurGLeuHisProThrSerGlnProLeuPheAlaPhe 828
QY 3024 GAATGGAGAGATCCAGGTACGGAGAACCGGGAGCTCACTGGAGCCCGACCTGCCCAA 3083
Db 829 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 848
QY 3084 GGGTTCAAGAACTCCCGACCATCTTTGACGAGCCCTACACAGAGACTGGCCCACTTC 3143
Db 849 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 868
QY 3144 AGGATCCAAACACCTCAGGTGACCTCTCCAGTACGTGATGACCTGCTTCTGGGGGA 3203
Db 869 ArgIleGlnHisProAspLeuLeuLeuGlnTyrValAspAspLeuLeuAlaAla 888
QY 3204 GCCACCAACAGACTGCTTAGAAGGACGACGACTACTGCTGGAATGTCTGACCTA 3263
Db 889 ThrSerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu 908
QY 3264 GGTACAGAGCTCTGCTAAGAGGCCCGCATGCTGACGGAGCGACGAGAGAGGTAACTACTTGGG 3323
Db 909 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 928
QY 3324 TACAGTTTGGGACCGGACCGCATGCTGACGGAGCGACGAGAGAACTGAGTCCAG 3383
Db 929 TyrLeuLeuLysGluGlnArgTrpLeuThrGluAlaArgLysGlnThrValMetGly 948
QY 3384 ATACCGGCCCAACACAGCAACAAATAGAGAGTTTGGGGACAGCTGGAATTTGC 3443
Db 949 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 968
QY 3444 AGACTGTGATCCCGGGTTGGACTTAGCTAGAGCCCTCTACCGCTAACCAAGAA 3503
Db 969 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 988
QY 3504 AAAGGGGAATTCCTCGGGCTCTCGAGCACCAAGAGCATTTGATGCTATCAAAAGGCC 3563
Db 989 GlyThrLeuPheAsnTrpGlyProAspGlnGlnLysAlaTyrGlnGluIleLysGlnAla 1008
QY 3564 CTGCTGAGCGCACTCTCTGCGCCCTCCCTGACGTAACCTTAACCTTTATGTG 3623
Db 1009 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 1028
QY 3624 GATGACGTAAGGAGTAGCCGGGAGTTTAAACCAACCTTAGGACCATGGAGAA 3683
Db 1029 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArgArg 1048
QY 3684 CTGTGCTCTACTGTCAAGAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGCTTG 3743
Db 1049 ProValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTrpProCysLeu 1068
QY 3744 AAGGCTATCGAGCTGCGCCATCTGCTGCAAGAGCGCTGACAAATTTGCTTTGGACAG 3803
Db 1069 ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln 1088
QY 3804 AATATACGTGTAATAGCCCCCATGATTTGGAGAACATCGTTGGGAGCCCCCAGACCGA 3863
Db 1089 ProLeuValIleLeuAlaProHisAlaValGluAlaValLysGlnProProAspArg 1108
QY 3864 TGGATGACCAACGCCCGCATGCCACTATCAAGCCTCTCTC ---ACAGAGGGGTC 3920
Db 1109 TrpLeuSerAsnAlaArgMetThrHisTyrGlnAlaLeuLeuAspThrAspArgVal 1128
QY 3921 ACGTTGCTCCACGACCGCTCTCAACCTGCGACTCTTCTGCTGAAGAGACTGATGAA 3980
Db 1129 GlnPheGlyProValValAlaLeuAsnProAlaThrLeuLeuPro ---LeuProGluGlu 1147
QY 3981 CAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCGCGAAGGACCTT 4040
Db 4041

1148 GlyLeuGlnHisAsnCysLeuAspIleLeuAlaGluAlaHisGlyThrArgProAspLeu 1167
QY 4041 ACAGACATACCGCTGACTGGAGAAGTCTAACTGTTCTACTACGGAACGACGATGTG 4100
Db 1168 ThrAspGlnProLeuProAspAlaAspHisThrTrpTyrThrAspGlySerLeuLeu 1187
QY 4101 GTGAAAGTAAAGAGTGGCTGGGCGCGGTGGACGGGACCCCGACGATCTGGGCC 4160
Db 1188 GlnGluGlyGlnArgLysAlaGlyAlaAlaValThrThrGluThrGluValIleTrpAla 1207
QY 4161 AGCAGCTCGCGGAGGAACCTTCAGCACAAAGCGCTGAGCTCATGGCCCTCACGCAAGCT 4220
Db 1208 LysAlaLeuProAlaGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAla 1227
QY 4221 TTGCGCTCGCGGAGGAAATCCATAAACATTATATACGACAGCAGGTATGCTTTGCG 4280
Db 1228 LeuLysMetAlaGluGlyLysLysLeuAsnValTyrThrAspSerArgTyrAlaPheAla 1247
QY 4281 ACTGCACAGTACATGGGCGCATCTATAAACAAAGGGGTGCTTACCTCAGCAGGAGG 4340
Db 1248 ThrAlaHisIleHisGlyGluIleTyrArgArgGlyLeuLeuThrSerGluGlyLys 1267
QY 4341 GAAATAAGAACAAAGGAAATTCCTAAGCCTATTAGAAGCCGTACATTTACCAAAAGG 4400
Db 1268 GluIleLysAsnLysAspGluIleLeuAlaLeuLysAlaLeuPheLeuProLysArg 1287
QY 4401 CTAGCTATTATACACTGTCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAAC 4460
Db 1288 LeuSerIleIleHisCysProGlyHisGlnLysGlyHisSerAlaGluAlaArgGlyAsn 1307
QY 4461 CAGATGGCTGACCGGGTTGCCAAGCAGCGCCAGCGGTGTTAAACCTTCTGCTCATATA 4520
Db 1308 ArgMetAlaAspGlnAlaAlaArgLysAlaAla ---IleThr 1320
QY 4521 GAAATGCCCAAGGCC ---CCAGAACCCAGA 4547
Db 1321 GluThrProAspThrSerThrLeuLeuIleGluAsnSerSerProTyrThrSerGluHis 1340
QY 4548 CGACGATACACCTAGACACTGGCAAGAGATAAAAGATAGACCCAG ---TTCTCTGAG 4604
Db 1341 PheHisTyrThrValThrAspLysAspLeuThrLysLeuGlyAlaIleTyrAspLys 1360
QY 4605 ACTCGGAAGGACCTGTATACCTCAGATGGGAAGAAATCTGCCCCCAACAAAGAGGG 4664
Db 1361 ThrLysLysTyrTrpValTyr ---GlnGlyLysProValMetProAspGlnPheThr 1378
QY 4665 TTAGATATGTCAACAGATACATCGTCTAACCCACTAGGAACCTAAACCTGCGACGAG 4724
Db 1379 PheGluLeuLeuAspPheLeuHisGlnLeuThrHisLeuSerPheSerLysMetLysAla 1398
QY 4725 TTGGTCAGAAC ---TCCCTTATCATGTTCTGAGGCTACCCAGAGTGGCTGACTCG 4778
Db 1399 LeuLeuGluArgSerHisSerProTyrTyrMetLeuAsnArgAspArgThrLeuLysAsn 1418
QY 4779 GTGGTCAACATGTGTGCTGCCCTGCCAGCTGGTTAATGCTTAATCTTCCAGATGCTCCA 4838
Db 1419 IleThrGluThrCysLysAlaCysAlaGlnValAlaSerLysSerAlaValLysGln 1438
QY 4839 GGAAGAGACTAAGGGAGAGCCAGCCAGCGCTCACTGGGAAGTGAAGTCTCACTGAGGTA 4898
Db 1439 GlyThrArgValArgGlyHisArgProGlyThrHisTrpGluIleAspPheThrGluIle 1458
QY 4899 AAGCGGCTAAATACGGAACAAATACCTATTGTTTGTAGACACCTTTTCAGGATG 4958
Db 1459 LysProGlyLeuTyrGlyTyrLysTyrLeuLeuValPheIleAspThrPheSerGlyTrp 1478
QY 4959 GTAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTTAAATAATACTGGAA 5018
Db 1479 IleGluAlaPheProThrLysLysGluThrAlaLysValValThrLysLysLeuLeuGlu 1498
QY 5019 GAAATTTTCCAAAGATTGGGAATACCTAAGGTAATAGGTGAGCAATCGTCAGGCTTT 5078
Db 1499 GluIlePheProArgPheGlyMetProGlnValLeuGlyThrAspAsnGlyProAlaPhe 1518

649	GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu	668
2547	TACTCTCCCTAGTAAAGCCTGATCAAAATATA---CAATTCTGGTTGGACAGTTTCCC	2603
669	HisGluThrSerIysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro	688
2604	CAAGCCTGGCGAGAAACCGCAGGAGTGGTGTGGCAAAGCAAGTTCCTCCACCAAGTTATT	2663
689	GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle	708
2664	CAACTGAAGCCAGTGCACACCACTGTCTCAGTCAGACAGTACCCTCTGAGTAAGAAGCT	2723
709	ProLeuIysAlaThrSerThrProValSerIleGlySerGlnTrpProMetSerGlnGluAla	728
2724	CAAGAAGCAATTGGCCGCGATGTCCAAAGATTAAATCCAAAGGCGATCTCTAGTTCCTGTC	2783
729	ArgLeuGlyIleIysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys	748
2784	CAATCTCCCTGGAACTACCTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGA	2843
749	GlnSerProTrpAsnThrProLeuLeuProValIysLysProGlyThrAsnAspTyrArg	768
2844	CCAGTACAGCACTTGAGAGAGTCAATAAACGGGTGCAGGATATACACCCAAAGTCCCG	2903
769	ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro	788
2904	AACCTTTATAACCTTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGATTGGAC	2963
789	AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp	808
2964	TTAAGGATGCCTTCTTCTCCCTGAGATTACACCCCTAGCCCAACCACTTTTGGCTTC	3023
809	LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe	828
3024	GAATGAGAGATCCAGGTACGGGAAGAAACCGGCAGCTCACTGGAGCCGACGCCCAA	3083
829	GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln	848
3084	GGGTTCAGAAATCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGGCCAACTTC	3143
849	GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe	868
3144	AGGATCCAAACCACTCAGGTGACCTCTCTCCAGTACGTGGATCACTGCTTCTGGCGGGA	3203
869	ArgIleGlnHisProAspLeuIleLeuLeuGlnTyr-ValAspAspLeuLeuAlaAa	888
3204	GCACCAACACAGCACTGCTTAGAAGGCACGAAGGCCTACTGCTGGAAATGTCTGACCTA	3263
889	Thr-SerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu	908
3264	GGCTACAGAGCCTCTGCTAAGAGGCCAGATTGTCAGGAGAGAGAGTAAACATACTTGGGG	3323
909	GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly	928
3324	TACAGTTTGGGACGGGCAGCGATGGCTGACGGAGCCACGGAAGAAAACCTGTAGTCCAG	3383
929	Tyr-LeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly	948
3384	ATACCGGCCCCAAACACACAGCCAAAACAAATGAGAGAGTTTTTGGGGACAGCTGATTTGC	3443
949	GlnProThrProLysThrProArgGlnLeuArgLupheLeuGlyThrAlaGlyPheCys	968
3444	AGACTGTGATCCCGGGTTTGGACCTTAGCAGCCCCACTCTACCCGCTAACCAAGAA	3503
969	ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr	988
3504	AAAGGGGAATTCTCTCGGGCTCTGTAGACCAAGAGGCATTTGATCTCATCAAAAAGGCC	3563
989	GlyThrLeuPheAsnTrpGlyProAspGlnGlnLysAlaTyrGlnGluIleLysGlnAla	1008
3564	CTGCTGAGGCACCTGCTCTGGCCCTCCCTGAGCTAACTAAACCCCTTTACCTTTATGTG	3623
1009	LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal	1028

Qy	3624	GATGAGCGTAAAGGAGTAGCCCGGGAGTTTTTAACCCAAAACCTTAGACCAATGGAGAAGA	3683
Db	1029	AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTyrArgArg	1048
Qy	3684	CTGTCCGCTACCTGTCAAAGAGCTCGATCTCTGTAGCCAGTGGTGGCCCATATGCGCTG	3743
Db	1049	ProValAlaTyrLeuSerLysLysLeuAspProValAlaGlyTyrProProCysLeu	1068
Qy	3744	AAGGCTATCGCAGCTGTGGCCATCTCGTCAAGGAGCGCTGACAAATGACTTTGGGACAG	3803
Db	1069	ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln	1088
Qy	3804	AATATAACTGTAATAGCCCCCATGTATGGAGAACATCGTTGGCAGCCCCCAGACCCGA	3863
Db	1089	ProLeuValIleLeuAlaProHisAlaValGluAlaLeuValLysGlnProProAspArg	1108
Qy	3864	TGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTC---ACAGAGAGGCTC	3920
Db	1109	TrpLeuSerAsnAlaArgMetThrHisTyrGlnAlaLeuLeuAspThrAspArgVal	1128
Qy	3921	ACGTTTCGCTCCACCAGCGCTCTCAACCCCTGCCACTCTTCGCTGAAGAGACTGATGAA	3980
Db	1129	GlnPheGlyProValValAlaLeuAsnProAlaThrLeuLeuPro---LeuProGluGlu	1147
Qy	3981	CCAGTACTCATGATTCGCATCAACTATTGATGGAGAGACTGGGTCCCGAAGGACCTT	4040
Db	1148	GlyLeuGlnHisAsnCysLeuAspIleLeuAlaGluAlaHisGlyThrArgProAspLeu	1167
Qy	4041	ACAGACATACCGCTAGCTGGGAGAGTCTTAACCTGGTTCACTGACGGAAGCACTATGTC	4100
Db	1168	ThrAspGlnProLeuProAspAlaAspHisThrTyrTyrThrAspGlySerSerLeuLeu	1187
Qy	4101	GTGGAAGTAAAGAGATGGCTGGGGCGCGGTGTGACGGGACCCGCACGACTCGGCC	4160
Db	1188	GlnGluGlyGlnArgLysAlaGlyAlaAlaValThrThrGluThrGluValIleTyrAla	1207
Qy	4161	AGCAGCTGCGGGAAGAACTTCAGCACAAAGGCTGAGCTCATGGCCCTCAGCGAAGCT	4220
Db	1208	LysAlaLeuProAlaGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAla	1227
Qy	4221	TTGCGGCTGGCGGAAGGAAATCCATAAACATTTATACGACAGCAGTATGCTTTGCG	4280
Db	1228	LeuLysMetAlaGluGlyLysLysLeuAsnValTyrThrAspSerArgTyrAlaPheAla	1247
Qy	4281	ACTGCACACGTACATGGGGCCATCTATAAACAAAGGGGGTGTCTACTCAGCAGGGAGG	4340
Db	1248	ThrAlaHisIleHisGlyGluIleTyrArgArgArgGlyLeuLeuThrSerGluGlyLys	1267
Qy	4341	GAATAAAGAACAAAGAGGAAATCTTAAGCCTATTAGAACGGTACATTTACCAAAAGG	4400
Db	1268	GluIleLysAsnLysAspGluIleLeuAlaLeuLeuLysAlaLeuPheLeuProLysArg	1287
Qy	4401	CTAGCTATTATACACTGCTCGGACATCCAGAAGCTCAATATCCAGAGGAAAC	4460
Db	1288	LeuSerIleIleHisCysProGlyHisGlnLysGlyHisSerAlaGluAlaArgGlyAsn	1307
Qy	4461	CAGATGCTGACCGGGTGGCAAGCAGCGACGCCAGGGGTGTAACCTTCTGCCCTATAATA	4520
Db	1308	ArgMetAlaAspGlnAlaAlaArgLysAlaAla-----IleThr	1320
Qy	4521	GAATGGCCCAAGCC-----CCAGAACCCAGA	4547
Db	1321	GluThrProAspThrSerThrLeuLeuIleGluAsnSerSerProTyrThrSerGluHis	1340
Qy	4548	CGACAGTACACCTTAGAAGACTGCGCAAGAGATAAAAAAGATAGACCAG---TTCTCTGAG	4604
Db	1341	PheHisTyrThrValThrAspIleLysAspLeuThrLysLeuGlyAlaIleTyrAspLys	1360
Qy	4605	ACTCCGGAAGGACCTCGTATACCTCAGATGGGAAGGAATCCTGCGCCCAACAAGAGGG	4664
Db	1361	ThrLysLysTyrTyrValTyr-----GlnGlyLysProValMetProAspGlnPheThr	1378

RESULT 4


```

US-09-075-272-4
; Sequence 4, Application US/09075272
; Patent No. 6136598
; GENERAL INFORMATION:
; APPLICANT: MILLER, A. DUSTY
; APPLICANT: WOLGAMOT, GREG
; APPLICANT: BOSHAM, LYNN
; TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL
; TITLE OF INVENTION: PACKAGING CELL LINES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,272
; FILING DATE: 08-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,140
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: 1438A-003710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-075-272-4

```

QY	2460	TCTTTTGGACAAGGGAACACGAGAAGTGTGTGCAAAATAACAAACCTATCATCTGTTGTGACC	2519
DB	101	 GlnPheThrSerGluGlyProGlnValSerTrpGlyLeuAlaProLeuAlaCysLeuVal	120
QY	2520	CTCCAAATTTAGATGACGAATATCGACTATCTCTCCCTAGTAAAGCCTGATCAAAATATA	2579
DB	121	 LeuSerThrGluGluGluTyArgLeuHisGluGlnProLysGlyAlaAlaProLeu	140
QY	2580	CAATTCCTGGTTGGAACAGTTTCCCAAGACCCCTGGGCAGAAAACCGCAGGATGGGTTTGGCA	2639
DB	141	 Asp-- -TrpValThrAlaPheProAsnValTrpAlaGluGlnAlaGlyMetGlyLeuAla	159
QY	2640	AAGCAAGTTTCCCCACAAAGTTATTCAACTGAAGCCAGTGCACACACAGTGTCTAGTCAGA	2699
DB	160	 LysGlnValProProValValValGluLeuLysAlaAspAlaThrProLysSerValArg	179
QY	2700	CAGTACCCCTTGACTAAAGAGCTCAAGAGGAATTCGGCGCATGTCCAAAGATTATC	2759
DB	180	 GlnTyProMetSerLysGluAlaLysGluGlyLeuArgProHisLeuArgLeuLeu	199
QY	2760	CAACAGGGCATCTCTAGTTCCTGTCCTCAATCTCCCTGGAAATCTCCCTGTCTACCGGTTAGA	2819
DB	200	AspGlnGlyLeuLeuValAlaCysGlnSerProTrpAsnThrProLeuLeuProValArg	219
QY	2820	AAGCCTGGGACTAATGACTATCGACACAGTACAGAGCTTGAGAGAGGTCAATAAACGGGTG	2879
DB	220	LysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgVal	239
QY	2880	CAGGATATACACCAACAGTCCCGAACCCCTTATAACCTCTTGTGTCTCTCCACACCCAA	2939
DB	240	LeuAspIleHisProThrValProAsnProTyrAsnLeuLeuSerSerLeuProProGlu	259
QY	2940	CGAGCTGGGTATACAGTATTGGACTTAAAGGATGCCCTTCTTCTGCTCGAGATTACACCCC	2999
DB	260	ArgThrTrpTyThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisPro	279
QY	3000	ACTAGCCAACCACTTTTGGCTTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAG	3059
DB	280	LysSerGlnLeuLeuPheAlaPheGluTrpArgAspProGluGlyGlnThrGlyGln	299
QY	3060	CTACCTGGACCCGACTGCCCGAAGGTTCAAGAACTCCCGACCATCTTTTGACGAAGCC	3119
DB	300	LeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAspGluAla	319
QY	3120	CTACACAGAGACTGGCCCACTTCAGGATCCACACCCCTCAGCTGACCTCTCCCTCAGTAC	3179
DB	320	LeuHisArgAspLeuAlaProPheArgAlaGlnAsnProGlnLeuThrLeuLeuGlnTyr	339
QY	3180	GTGATGACCTGTCTTCGGCGGAGCCACCAACAGGACTGCTCTAGAAGCAGCAAGGCA	3239
DB	340	ValAspAspLeuLeuIleAlaAlaSerLysGluLeuCysGlnGlnGlyThrGluArg	359
QY	3240	CTACTGTGGAATGTCTGACCTAGGCTACAGAGCCTCTGCTTAAGAAGGCCACAGATTGC	3299
DB	360	LeuLeuThrGluLeuGlyAsnLeuGlyTyrArgValSerAlaLysLysAlaGlnIleCys	379
QY	3300	AGGAGAGCTAACATCTTGGGTACAGTTTCGGGACGGGCAGGATGGCTCAGCGGAG	3359
DB	380	GlnThrGluValIleTyrLeuGlyTyrThrLeuArgGlyGlyLysArgTrpLeuThrGlu	399
QY	3360	GCACGAAGAAAACCTGTAGTCCAGATACCGGCCCCCAACACAGCCAAACAAATAGAGAG	3419
DB	400	AlaArgLysLysThrValMetMetIleProProProThrThrProArgGlnValArgGlu	419
QY	3420	TTTTTGGGACAGCTGGATTTTCAGACTGTGGATCCCGGGGTTTCGACCTTTAGCAGCC	3479
DB	420	PheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaAla	439
QY	3480	CCACTTACCCGCTAACCAAGAAAAAGGGGAATTTCTCTGGGCTCCTGAGCACCAGAG	3539
DB	440	ProLeuTyProLeuThrArgGluGlyIleProPheGluTrpLysGluHisGlnArg	459

Db 1168 ThrAspAsnProLeuLysLeuArgLeuLeuArgSerPro 1181

RESULT 5

US-08-929-967-8
Sequence 8, Application US/08929967
Patent No. 5891637
GENERAL INFORMATION:

APPLICANT: Rupert, Siegfried J.W.
TITLE OF INVENTION: Construction of Full-Length cDNA Libraries
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/929,967

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Ginger R.

REGISTRATION NUMBER: 33,055

REFERENCE/DOCKET NUMBER: P1035R1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-3216

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1079 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-929-967-8

Alignment Scores:

Pred. No.:	1.18e-299	Length:	1079
Score:	3600.00	Matches:	679
Percent Similarity:	76.7%	Conservative:	156
Best Local Similarity:	62.4%	Mismatches:	222
Query Match:	24.6%	Indels:	32
DB:	1	Gaps:	11

US-10-723-552-3 (1-8132) x US-08-929-967-8 (1-1079)

QY	2517	ACCCTCCAAATTAGATGACGAATATCGACTATATCTCTCCCTAGTAAAGCCTGATCAAAAT	2576
Db	1	ThrLeuAsnIleGluAspGluHisArgLeuHisGluThrSerLysGluProAspValSer	20
QY	2577	ATA---CAATTCGTGGTGAACAGATTTCCTCCAGCTGGGCGAGAAACCCAGGGATGGGT	2633
Db	21	LeuGlySerThrTriLeuSerAspPheProGlnAlaTriPalaGluThrGlyMetGly	40
QY	2634	TTGGCAAGCAAGTTCCTCCCAAGTTATTCACTGAAGCCAGTGCCACACAGTGCA	2693
Db	41	LeuAlaValArgGlnAlaProLeuIleIleProLeuLysAlaThrSerThrProValSer	60
QY	2694	GTGACAGCTACCTCTGAGTAAAGAGCTCAAGAGGAATTCGGCGCATGTCCAAAGA	2753
Db	61	IleLysGlnTyrProMetSerGlnGluAlaArgLeuGlyIleLysProHisIleGlnArg	80
QY	2754	TTAATCCAAAGGGCATCTAGTCTCTGCTCAATCTCTCCCTGGAATATCTCCCTGCTACCG	2813
Db	81	LeuLeuAspGlnGlyIleLeuValProCysGlnSerProTriPsnThrProLeuLeuPro	100
QY	2814	GTTAGAAAGCTGGGACATATGACTATCGACCAAGTACAGGACTTGAGAGAGTCAATAAA	2873
Db	101	ValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLys	120

QY	2874	CGGGTCAGGATATACACCCAAAGTCCCGAACCTTATAACCTTCTGTGTCTCTCCCA	2933
Db	121	ArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeuPro	140
QY	2934	CCCCAAGGAGCTGGTATACAGTATTGGACTTAAGGATGCTCTTCTGCTCGAGATTAA	2993
Db	141	ProSerHisGlnTriPtyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeu	160
QY	2994	CACCCCACTAGCCAAACCTCTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGAAC	3053
Db	161	HisProThrSerGlnProLeuPheAlaPheGluTriPargAspProGluMetGlyIleSer	180
QY	3054	GGGCAGCTCACCTGGACCCGACTGCCCAAGGTTCAAGAATCTCCCGACCATCTTTGAC	3113
Db	181	GlyGlnLeuThrTriPtyrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAsp	200
QY	3114	GAAGCCCTACACAGAGACTGGCCAACTTCAGATCCACACCTCAGGTGACCTCCCTC	3173
Db	201	GluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuIleLeuLeu	220
QY	3174	CAGTACGTGATGACCTGCTCTGGCGGAGCCACCAACAGACACTGCTTAGAAGGCACG	3233
Db	221	GlnTyrValAspAspLeuLeuAlaThrSerGluLeuAspCysGlnGlnGlyThr	240
QY	3234	AAGCAGCTACTGCTGGAAATTGCTGACCTAGGCTACAGAGCTCTGCTTAAGAGGCCAG	3293
Db	241	ArgAlaLeuLeuGlnThrLeuGlyAsnLeuGlyTyrArgAlaSerAlaLysAlaGln	260
QY	3294	ATTGCGAGGAGAGAGTAACTACTTTGGGTTACAGTTTCGGGAGCGGGAGGATGGCTG	3353
Db	261	IleCysGlnLysGlnValLysTyrLeuGlyTyrLeuLysGluGlyGlnArgTriLeu	280
QY	3354	ACGGAGGCGCAGGAAGAACTAGTCCAGATACCGGCCCAACACACGCCCAACAAATG	3413
Db	281	ThrGluAlaArgLysGluThrValMetGlyGlnProThrProLysThrProArgGlnLeu	300
QY	3414	AGAGAGTTTTTGGGACAGCTGGATTTTGCAGACTGTGATCCCGGGTTTGGCGACTTA	3473
Db	301	ArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTriPleProGlyPheAlaGluMet	320
QY	3474	GCAGCCCACTCTACCGCTAACCAAGAAAGGGAATTCCTCTGGGCTCTGAGCAC	3533
Db	321	AlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPheAsnTriPglyProAspGln	340
QY	3534	CAGAAGGCATTTGATGCTATCAAAAGCCCTGAGCGCACCTGCTGCGCCCTCCCT	3593
Db	341	GlnLysAlaTyrGlnGluLysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeuPro	360
QY	3594	GACGTAACTAAACCTTTTACCCTTTTATGTGATGAGCGTAAAGGAGTAGCCCGGGAGTT	3653
Db	361	AspLeuThrLysProPheGluLeuPheValAspGluLysGlnGlyTyrAlaLysGlyVal	380
QY	3654	TTAACCCAAACCTTAGGACCATGGAAGAGACTGTGCGCTACTGTCAAAAGAGCTCGAT	3713
Db	381	LeuThrGlnLysLeuGlyProTriPargArgProValAlaTyrLeuSerLysLysLeuAsp	400
QY	3714	CCTGTAGCCAGTGGTGGCCCATATGCTCGAAGGCTATCGCAGCTGTGGCCATCTGGTC	3773
Db	401	ProValAlaAlaGlyTriPProProCysLeuArgMetValAlaAlaIleAlaValLeuThr	420
QY	3774	AAGGACGTGACAAATTGACTTTGGGACAGATAATAACTGTAAATAGCCCCCATGCTATG	3833
Db	421	LysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAlaVal	440
QY	3834	GAGAACTCGTTGGCAGCCCCCAGACCGATGATGACAAAGCCCGCATGACCCACTAT	3893
Db	441	GluAlaLeuValLysGlnProProAspArgTriLeuSerAsnAlaArgMetThrHisTyr	460
QY	3894	CAAGCCCTGCTCTC---ACAGAGAGGCTCAGTTTCGCTCCACACGCGCTCTCAACCT	3950
Db	461	GlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProValAlaLeuAsnPro	480

```
QY 3951 GCCACTCTCTGCTGAAGAGACTGTGAACCACTGACTCATGATGGCCATCAACTATTG 4010
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 AlaThrLeuLeuPro----LeuProGluGluGluGluGlnHisAsnCysLeuAspIleLeu 499
QY 4011 ATTGAGGAGACTGGGGTCCCAAGGACTTTACAGACATACCTCGCTGACTGGAGAAGTGCTA 4070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 AlaGluAlaHisGlyThrArgProAspLeuThrAspGlnProLeuProAspAlaAspHis 519
QY 4071 ACCTGGTTCACTACGGAACAGCTATGCTGTGAAGGATGAAGGATGCTGGGGGGCGG 4130
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
520 ThrTrpTrpThrAspGlySerSerLeuLeuGlnGluGlyGlnArgLysAlaGlyAlaAla 539
QY 4131 GTGGTGGACGGGACCCGACAGATCTGGGCCAGAGCCTCGCCGGAAGAACTTCAGACAAA 4190
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
540 ValThrThrGluThrGluValIleTrpAlaLeuAlaLeuProAlaGlyThrSerAlaGln 559
QY 4191 AAGCTGAGTCAATGGCCCTCAGCAAGCTTTGGCTGGCCGCAAGGAAATCCATAAAC 4250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
560 ArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysLysLeuAsn 579
QY 4251 ATTATACGACAGCAGGATGCTTTGCGACTGCACAGTACATGGGGCCATCTATAA 4310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
580 ValTrpThrAspSerArgTrpAlaPheAlaThrAlaHisIleHisGlyGluIleTrpArg 599
QY 4311 CAAAGGGGGTGTACTCTCAGCAGGGAGGAAATAAGAAACAAAGAGGAAATCTAAGC 4370
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 ArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAspGluIleLeuAla 619
QY 4371 CTATTAGAAGCCGTACATTTACCAAAAGGCTAGTATTATATACATGCTCTCGACATCAG 4430
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
620 LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHisGln 639
QY 4431 AAGCTAAAGATCTATATCCAGAGAAACAGATGCTGACCGGTGGCCAGACAGGCA 4490
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
640 LysGlyHisSerAlaGluAlaArgGlyAsnArgMetAlaAspGlnAlaAlaArgLysAla 659
QY 4491 GCCAGGGTGTAACTTCTGCTCTATATAAGAAATGCCAAAGCC----- 4535
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
660 Ala-----IleThrGluThrProAspThrSerThrLeuLeuIle 672
QY 4536 -----CCAGAACCCAGACGACGTACACCTAGAGAGACTGGCAAGAG 4577
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
673 GluAsnSerSerProTrpThrSerGluHisPheHisTrpValThrAspIleLysAsp 692
QY 4578 ATAAAGAGATAGACAG-----TTCTGTGACACTCCGGAAGGACCTCTATACCTCAGAT 4634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
693 LeuThrLysLeuGlyAlaIleTrpAspLysThrLysLysTrpValTrp-----Gln 710
QY 4635 GGAAGGAAATCTGCCCCCAAGAGGGTTAGAAATATGTCCAAACAGATACATCGTCTA 4694
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
711 GlyLysProValMetProAspGlnPheThrPheGluLeuLeuAspPheLeuHisGlnLeu 730
QY 4695 ACCCACTTAGGAACCTAACACCTGCACAGTTGGTGCAGAAC-----TCCCTTTATCAT 4748
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
731 ThrHisLeuSerPheSerLysMetLysAlaLeuLeuGluArgSerHisSerProTrpTrp 750
QY 4749 GTTCTGAGGCTACAGAGAGTGGCTGACTCGGTGTGCTCAAAATTTGTGTGCCCTGCCAGCTG 4808
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 MetLeuAsnArgAspArgThrLeuLysAsnIleThrGluThrCysLysAlaCysAlaGln 770
QY 4809 GTTAATGCTAATCTCCAGAACTCCAGAGGAGAGACTAAGGGGAAGCCACCCAGGC 4868
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
771 ValAsnAlaSerLysSerAlaValLysGlnGlyThrArgValArgGlyHisArgProGly 790
QY 4869 GCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGCTAAATACGGAACAAATACCTA 4928
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
791 ThrHisTrpGluIleAspPheThrGluIleLysProGlyLeuTrpGlyTrpLysTrpLeu 810
QY 4929 TTGGTTTTTGTAGACACTTTTCAGGATGGGTAGAGCTTATCTCTACTAAGAAAGACT 4988
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
811 LeuValPheIleAspThrPheSerGlyTrpIleGluAlaPheProThrLysLysGluThr 830
QY 4989 TCAACCGTGGTGGCTAAAAAATACTGGAAGAAATTTTCCAGATTTGGGAATACCTAAG 5048
```

```
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
831 AlaLysValValThrLysLysLeuLeuGluGluPheProArgPheGlyMetProGln 850
QY 5049 GTAATAGGTGACACAATGCTCAGCTTTTGTGGCCAGGTAAGTCAAGGACTGGCCAAAG 5108
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 ValLeuGlyThrAspAsnGlyProAlaPheValSerLysValSerGlnThrValAlaAsp 870
QY 5109 ATATTGGGATTCATTGGAACTGTGATTGTGCATACAGACCCCAAGCTCAGACAGGTA 5168
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
871 LeuLeuGlyIleAspTrpLysLeuHisCysAlaTrpArgProGlnSerGlyGlnVal 890
QY 5169 GAGAGATGAATAGAACCACTTAAAGAGACCTTTACTAAATTTGACCCGGAGACTGCGTT 5228
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
891 GluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrLeuAlaThrGlySer 910
QY 5229 AATGATTGATAGTCTCTCTGCGCTTTGCTTTTGGTGTAGGAACACCCCTGGACAG 5288
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
911 ArgAspTrpValLeuLeuLeuProLeuAlaLeuTrpArgAlaArgAsnThrProGlyPro 930
QY 5289 TTTGGGCTGACCCCTATGAATTTACTCTACGGGGGACCCCCCAATTTGGTAGAAATTGCT 5348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
931 HisGlyLeuThrProTrpGluIleLeuTrpGlyAlaProProLeuValAsnPhePro 950
QY 5349 TCTGTACATAGTGTGCTGCTGCTTTCCAGCCCTTTTCTCTAGGCTCAAGGCACTT 5408
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
951 AspProAspMetThrArgValThrAsnSerProSerLeuGlnAlaHisLeuGlnAlaLeu 970
QY 5409 GAGTGGGTGAGACAAAGCGGTGGAGGCACTCCGGGAGGCTTACTCAGGAGGAGAGAC 5468
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
971 TyrLeuValGlnHisGluValTrpArgProLeuAlaAlaIleArgGlnGluGlnLeuAsp 990
QY 5469 TTGCAG--ATCCCATCGTTTCCAAAGTGGGAGATTCTAGCTTACGCTAGAGCCCAACGT 5525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
991 ArgProValValProHisProTrpArgValGlyAspThrValTrpValArgArgHisGln 1010
QY 5526 GCAGGAAACCTCGAGACTCGGTGGAGGCGCCCTTATCTCGTACTTTTACCAACCAACG 5585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1011 ThrLysAsnLeuGluProArgTrpLysGlyProTrpThrValLeuLeuThrThrProThr 1030
QY 5586 GCTGTGAAGTGAAGGAATCTCCACTGATCCATCCATCCATCCATCCATCCATCCATCC 5645
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1031 AlaLeuLysValAspGlyIleAlaAlaTrpIleHisAlaAlaHisValLysAlaAlaAsp 1050
QY 5646 CCT-----CCCGATTTCG-----GGGTGAAAGCCGAAAGACTGAAATCCCTT 5690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1051 ProGlyGlyGlyProSerSerArgLeuThrTrpArgValGlnArgSerGlnAsnProLeu 1070
QY 5691 AAGCTTTCGCTCCATCGCGTGGTTCCT 5717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1071 LysIleArgLeuThrArgGluAlaPro 1079
```

RESULT 6

```
US-09-376-784-6
Sequence 6, Application US/09376781
Patent No. 6261806
GENERAL INFORMATION:
APPLICANT: Bacterio Papias T.
APPLICANT: Patience, Clive
APPLICANT: Anderson, Goran K.
TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
Patent No. 6261806
TITLE OF INVENTION: Use
FILE REFERENCE: 61750-267
CURRENT APPLICATION NUMBER: US/09/376,781
CURRENT FILING DATE: 1999-08-18
EARLIER APPLICATION NUMBER: 60/097,015
EARLIER FILING DATE: 1998-08-18
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 638
TYPE: PRN
ORGANISM: Artificial Sequence
```

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:PERV-C
; OTHER INFORMATION: polypeptide sequence taken from GenBank Accession
; OTHER INFORMATION: No. 6261806 AF038600 for comparison.
US-09-376-781-6

Alignment Scores:

Pred. No.: 8.95e-285 Length: 638
Score: 3425.00 Matches: 637
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 23.4% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3 (1-8132) x US-09-376-781-6 (1-638)

QY 5620 ATGCATCCCAAGCTTAAACGGCGCCACCTCCCGATTCGGGGTGGAAAGCGGAAGACTG 5679
DB 1 MeHisProThrLeuAsnArgHisLeuProIleArgGlyGlyProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATAAACCCTCTCAG 5739
DB 21 LysIleProLeuSerPheAlaSerIleAlaTirPheLeuThrLeuSerIleThrSerGln 40
QY 5740 ACTAATGGTATGGCATAGGAGACAGCTGAACCTCCCATAAACCCTTATCTCTCAGCTGG 5799
DB 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTir 60
QY 5800 TTAATTACTACATCCGSCACAGGTATTAAATCAACAACACTCAAGGGAGGCTCCTTTA 5859
DB 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyAlaProLeu 80
QY 5860 GGAACCTGGTGGCTGATCTATACGTTGGCTGCAGATCAGTTATTCCTAGTCTGACCTCA 5919
DB 81 GlyThrTirProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 5920 CCCCAGATATCTCATGCTCAGGATTTTATGTTTGGCCAGGACCAACAATAATGGA 5979
DB 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 5980 AAACATTGGGAATCCAGAGATTTCTTTTGAACAAATGGAACGTGTACCTCTAAT 6039
DB 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTirPasnCysValThrSerAsn 140
QY 6040 GATGGATATTGGAATGGCAACCTCTCAGCAGGATAGGTAAAGTTTCTTATATGCAAC 6099
DB 141 AspGlyTyrTirPlystrProThrSerGlnInAspArgValSerPheSerTyrValAsn 160
QY 6100 ACCTATACAGCTCTGGACAATTAATTACCTGACCTGGATTAGAACTGGAAGCCCAAG 6159
DB 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTirPileArgThrGlySerProLys 180
QY 6160 TGCTCTCCTCAGACTAGATTACCTAATAATAGTTTCACTGAGAAAGGAAGAAACAGAA 6219
DB 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGlnLysGlyLysGlnGlu 200
QY 6220 AATATCTTAAATGGGTAAATGGTATGCTTGGGGAATGGTATATTATGAGGCTCGGT 6279
DB 201 AsnIleLeuLysTirPalleAsnGlyMetSerTirPlyMetValTyrTyrGlySerGly 220
QY 6280 AAACAACAGGCTCCATTTCAATATTTCGCTCAAAATAAACAGCTGAGCTCCCAATG 6339
DB 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 6340 GCTATAGGACCAATACGCTTGACGGGTCAAGACCCCAAGCCCAAGGCCAGGACCA 6399
DB 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProProThrGlnGlyProGlyPro 260
QY 6400 TCCTCTAAACATAACTTCTGGATCAGACCCCACTGAGTCTAAACAGACGACTAAATGGGG 6459
DB 261 SerSerAsnIleThrSerGlySerAspProThrGlnSerSerSerThrThrLysMetGly 280
QY 6460 GCAAAACTTTTAGCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAG 6519

RESULT 7

US-09-111-085-2

DB 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 6520 GCTACCTCTTCTTGGCTATGCTTACGCTTCCGGGCCACCTTACTATGAGGAATGGCT 6579
DB 301 AlaThrSerSerCysTirPleuCysLeuAlaSerGlyProProTyrTyrGluGlyMetAla 320
QY 6580 AGAAGAGGGAATTCATATGTGACAAAGAACATAGAGACCAATGCACATGGGATCCCAA 6639
DB 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTirPlySerGln 340
QY 6640 AATAAGCTTACCTTACTGAGGTTCTTGAAAAGGACCTGCATAGAAAAGTTCCCCCA 6699
DB 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 6700 TCCACCAACACACTTGTAAACACACTGAAGCCTTTAATCARACCTCTGAGAGTCAATAT 6759
DB 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 6760 CTGGTACCTGGTTATGACAGGTGGTGGCATGTAAATACTGGATTAAACCCCTTGTGTTTC 6819
DB 381 LeuValProGlyTyrAspArgTirPAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 6820 ACCTGGTTTTAAACCAAACTAAAGATTTTGGTATTTGGTCCAAATTTGTTCCCCAGTG 6879
DB 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 6880 TATTACTATCCCGAAAAGCAATCCTTGATGAATATGACTACAGAAATATCTCACAAAAG 6939
DB 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 6940 AGAAGAACCATATCTCTGACACTGTCTGATGCTCGGACTTGGAGTGGCAGCAGGTGA 6999
DB 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaGlyVal 460
QY 7000 GGAACGAGAACAGCTGCCCTGTCCAGGACACACAGCAGCTAGAAAACAGACTTAGTAAC 7059
DB 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 7060 CTACATCGAATTTGTAACAGAAGATCTCCAAGCCCTAGAAAATCTCTCAGTACCTGGAG 7119
DB 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 7120 GAATCCCTTAACCTCTTATCTGAAGTAGTCTTACAGAATAGAGAGGGTTAGATTATTA 7179
DB 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 7180 TTTCTAAAAGAGGAGGATTTATGTAGCTTGAAGGAGAAATGCTGTTTTATGTGGAT 7239
DB 521 PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
QY 7240 CATTCAAGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAAGGTTGGAGAGCCCTCGA 7299
DB 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 7300 AGGAAAAGGAACTACTCAAGGTTGGTTGAGGGATGCTTCAACAGGCTCTCTTGGTTG 7359
DB 561 ArgGluLysGluThrThrGlnGlyTirPheGluGlyTirPheAsnArgSerLeuTirPLeu 580
QY 7360 GCTACCTACTTCTTCTGCTTTAAACAGACCTTAATAGTCTCTCTCTCTCTTACTCAGATT 7419
DB 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 7420 GGGCCATGTATTATTAACAAGTTAATGCTTCTCATTAGAGAAGTAATAGTCAGTCCAG 7479
DB 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 7480 ATCATGGTACTTAGACAACAGTACCAAGCCCTCTAGCAGGAAGCTGGCCGC 7533
DB 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638

```
; Sequence 2, Application US/091111085
; Patent No. 6100034
; GENERAL INFORMATION:
; APPLICANT: Stoye, Jonathan P
; APPLICANT: Weiss, Robin A
; TITLE OF INVENTION: Detection of retroviral subtypes based upon envelope
; FILE OF INVENTION: specific sequences
; FILE REFERENCE: 4238/75168
; CURRENT APPLICATION NUMBER: US/09/111,085
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: GB 9710154.7
; EARLIER FILING DATE: 1997-05-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Porcine retrovirus
US-09-111-085-2

Alignment Scores:
Pred. No.:      8.18e-237      Length:      660
Score:          2866.50       Matches:    537
Percent Similarity: 88.6%     Conservative: 39
Best Local Similarity: 82.6%   Mismatches: 53
Query Match:     19.6%       Indels:     21
DB:              2           Gaps:       6

US-10-723-552-3 (1-8132) x US-09-111-085-2 (1-660)
QY 5620 ATGCATCCCAACGTAAACCGCGCCACCTCCCATTCGGGGTGGAAAGCCGAAAGACTG 5679
DB 1 MetHisProThrLeuSerArgArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGCTGCTTCTTACTCTGCAATAAATCTCTCAG 5739
DB 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 5740 ACTAATGGTATGGCAGATAGAGACAGCTGAACCTCCCATTAACCCCTTATCTCTCAGTGG 5799
DB 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 5800 TTAATTAATCACTCCGCGCAGAGTATTAATATCAACAACACTCAAGGGGAGGCTCTTTTA 5859
DB 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyAlaProLeu 80
QY 5860 GGAACCTGGTGGCTGATATACGTTTGCTCGATCAGTCACTTATCTAGTCTG----- 5913
DB 81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 5914 ----ACCTCACCACCATATCCCTCCATGCTCAGGATTTTATGTTGCCAGGACCAACCA 5970
DB 101 GlnAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
QY 5971 AATAATGGAAACATTCGCGAAATCCAGAGATTTCTTTTGTAAACAATGGAACTGTGTA 6030
DB 121 AsnAsnGluIuTyrCysGlyAsnProGlnAspPheCysLysGlnTrpSerCysIle 140
QY 6031 ACCTCTAATGATGATATTCGAAATGGCCAACTCTCAGCAGGATAGGGTAAGTTTCT 6090
DB 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY 6091 TATGTCACACCTATACCACTCTGGCAATTAATTAAC-----CTGACC 6135
DB 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY 6136 TGG-----ATTAGAACTGGAAAGCCCAAGTCTCTCTCTCCAGC 6174
DB 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 6175 CTAGATTACATAAAATGATTTTCACTGAGAAAGGAAAAACAAGAAATATCTCTAAATGG 6234
DB 201 LeuAspTyrLeuLysIleSerPheThrThrGluLysGlyLysGlnGlnAsnIleGlnLysTrp 220
```

```
QY 6235 GTAAATGGTATGCTTTGGGAATGGTATATTATGGAGGCTCGGGTAAACAACACAGGCTCC 6294
DB ValAsnGlyIleSerTrpGlyIleValTyrTyrGlySerGlyArgLysLysGlySer 240
QY 6295 ATTCTAACTATTGCGCTCAAAATA----AACAGCTGGAGCTCCAATGGCTATAGGACCA 6351
DB ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProValAlaIleGlyPro 260
QY 6352 AATACGGTCTTGACGGGTCAAGACCCCAACCCAA-----GGACAGGACCA 6399
DB AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGluArgProSerProAsnPro 280
QY 6400 TCCTCT---AACATACTTCTGGATCAGACCCCACTAGCTCTAACAGCAGCTATAAATG 6456
DB SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY 6457 GGGGCAAAACCTTTTAACTCCATCCAGGAGCTTTTCAAGCTCTTAACCTCCACAGCTCCA 6516
DB GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY 6517 GAGGCTACCTCTTCTTGGTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGAATG 6576
DB GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMet 340
QY 6577 GCTAGAAGAGGGAAATTCATATGTGACAAAGAACATAGAGACCAATGCACATGGGATCC 6636
DB AlaArgGlyGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 6637 CAAAATAAGCTTACCCCTTACTGAGGTTTCTGGAAGAGGCACCTGCATAGAAAGTTTCCC 6696
DB GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380
QY 6697 CATTCCCAACACACCTTTGTAAACACACTGAAGCCCTTTAATCAACCTCTGAGAGTCAA 6756
DB ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 6757 TATCTGCTACCTGGTTATGACAGTGTGGGCACTGTAATACTGATTAACCCCTTGTTGTT 6816
DB TyrLeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 6817 TCCACCTCTGGTTTTTAAACCAAACTAAAGATTTTTTGCAATTATGTCCTCAAAATTTGCCCGA 6876
DB SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 6877 GTGTATTACTATCCCAAAAAGCAATCTTGTATGATATATGACTACAGAAATCATCCACAA 6936
DB ValTyrTyrTyrProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460
QY 6937 AAGAGAGAACCACATATCTCTGACACTTGTGTCATGCTCGGACTTGGAGTGGCAGCAGGT 6996
DB LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGly 480
QY 6997 GTAGGAACAGGAACACGCTGCCCTTGTGTCAGGGACCAACAGCAGCTAGAAAACAGGACTTAGT 7056
DB ValGlyThrGlyThrAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
QY 7057 AACCTACATCGAATTTGACAGAAATCTCCAAAGCCCTAGAAAATCTGTCAGTAACCTG 7116
DB AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 7117 GAGGAATCCCTAACCTCTTATCTGAAGTAGTCTCTACAGAAATAGAAGAGGTTAGATTTTA 7176
DB GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 7177 TTATTTCTAAAGAGGAGGATTTATGCTAGCCCTCAAGAGGAATGCTGTTTTTATGTG 7236
DB LeuPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrVal 560
QY 7237 GATCATTCAGGGCCATCAGAGACTCCATGACAAAGCTTAGAAAAGTTGGAGACGCT 7296
DB AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 580
```



```
QY 7297 CGAAGGGAAGAAACTACTCAAGGGTGGTTTGGAGGATGGTTCAACAGGCTCTTTGG 7356
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 7357 TTGGCTACCTACTTTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTGTTACTCACA 7416
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 620
QY 7417 GTTGGCCCATGTAATTAACTAAGTTAAATTCCTTCATTAGACGAAATAAGTCAGTC 7476
Db 621 ValGlyProCysLeuLeuLeuAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
QY 7477 CAGATCATGTACTAGACACAGTACCAC 7506
Db 641 GlnLeuMetValLeuArgGlnGlnTrpGln 650

RESULT 8
US-09-376-781-5
; Sequence 5, Application US/09376781
; Patent No. 6261806
; GENERAL INFORMATION:
; APPLICANT: Banerjee, Papija T.
; APPLICANT: Patience, Clive
; APPLICANT: Anderson, Goran K.
; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
; Patent No. 6261806
; TITLE OF INVENTION: Use
; FILE REFERENCE: 61750-267
; CURRENT APPLICATION NUMBER: US/09/376, 781
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/097, 015
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PERV-A
; OTHER INFORMATION: polypeptide sequence taken from GenBank Accession
; OTHER INFORMATION: No. 6261806 Y12238 for comparison.
US-09-376-781-5

Alignment Scores:
Pred. No.: 8,18e-237 Length: 660
Score: 2866.50 Matches: 537
Percent Similarity: 88.6% Conservatives: 39
Best Local Similarity: 82.6% Mismatches: 53
Query Match: 19.6% Indels: 21
DB: 2 Gaps: 6

US-10-723-552-3 (1-8132) x US-09-376-781-5 (1-660)
QY 5620 ATGCATCCCACTTAAACCGGCGCCACTCCCGATTCGGGTGGAAAGCCGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCCCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 5740 ACTAATGGTATGGCATAGGACAGCCTGAACTCCCAATAACCTTATCTCTCACCTGG 5799
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 5800 TTAATTACTGCTCCGGCAGAGGTATTAATATCAACACACTCAAGGGAGGCTCCTTTA 5859
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 5860 GGAACCTGGTGCCTGATCTATACGTTTGCCTCAGATCAGTATTCTTAGTCTG 5913
Db 81 GlyThrTrpProGluLeuTrpValCysLeuArgSerValIleProGlyLeuAsnAsp 100
```

```
QY 5914 ---ACCTCACCCCGATATCTCTCCATGCTCAGGATTTATGTTTGGCCAGCACCA 5970
Db 101 GlnAlaThrProProAspValLeuArgAlaTrpGlyPheTrpValCysProGlyProPro 120
QY 5971 AATAATGGAATAACATTCGCGAAATCCAGAGATTTCTTTTGTAAACAATAGGAAGTGTGA 6030
Db 121 AsnAsnGluGluTrpCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysIle 140
QY 6031 ACCTCTAATGATGATATGGAATGCGCAACTCTCAGCAGGATAGGAGTAACTTTTCT 6090
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTrpSer 160
QY 6091 TATGTCAACACCTATACAGCTCTCGACAATTTAATTAC-----CTGACC 6135
Db 161 PheValAsnAsnProThrSerTrpAsnGlnPheAsnTrpGlyHisGlyArgTrpLysAsp 180
QY 6136 TGG-----ATTAGAACTGGAGGCCCCAGTCTCTCTCTCAGAC 6174
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 6175 CTAGATTACCTAAATAAGTTTCTAGAGAAAGGAAACAAGAAAATATCTTAAATGG 6234
Db 201 LeuAspTrpLeuLysIleSerPheThrGluLysGlyLysGlnGlnIleGlnLysTrp 220
QY 6235 GTAAATGTTATGTTGGGAAATGTTATATTGAGGCTCGGTAAACAACACAGGCTCC 6294
Db 221 ValAsnGlyIleSerTrpGlyIleValTrpGlyGlySerGlyArgLysLysGlySer 240
QY 6295 ATTCTAACTATTTCGGCTCAAAATA---AACCAAGTGGAGCTCCCAATGGCTATAGAACCA 6351
Db 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProProValAlaIleGlyPro 260
QY 6352 RATAAGGCTTCGCGGTCAAGAGACCCCAACCCAA-----CGACCCAGCACCA 6399
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGluArgProSerProAsnPro 280
QY 6400 TCCTCT---AACATAACTTCTCGATCAGACCCCACTGAGCTTAACACAGCAGCTAAATG 6456
Db 281 SerAspTrpAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY 6457 GGGGCAAACTTTTATGCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCAGACTCCA 6516
Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY 6517 GAGGCTACTCTTCTTGTGGCTATGCTTAGCTTCGGGCCCACTTACTATGAAGAAATG 6576
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTrpTrpGlyMet 340
QY 6577 GCTAGAAGAGGGGAAATTCATGTGACAAAAGAACATAGAGACCAATGCACATGGGGATCC 6636
Db 341 AlaArgGlyGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 6637 CAAAATAAGCTTACCTTACTGAGTTTCTGGAAGAGCACCTGCATAGGAAGGTTCCC 6696
Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380
QY 6697 CCATCCCAACACACTTTGTAAACCACTGAGCTTTTAACTCAACCTCTGAGAGTCAA 6756
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 6757 TATCTGGTACTGTTTATGACAGGTGGTGGGATGTAATCTACGATTAAACCTCTGTGT 6816
Db 401 TyrLeuValProGlyTrpAspArgTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 6817 TCCACCTGGTGTTTTAAACCAACTAAAGATTTTTCATTTATGGTCCAAATGTTTCCCCGA 6876
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 6877 GTGATTTACTATCCCGAAAGCAATCTCTGATGATATGATCAGATAAATCATCGACAA 6936
Db 441 ValTrpTrpProGluLysAlaValLeuAspGluTrpAspTrpArgTrpArgTrpAsnArgPro 460
```

```
QY 6937 AAGAGAAACCCATATCTCTGACACTTGTGTGATGCTCGACTTGGAGTGGCAGCAGGT 6996
Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
QY 6997 GTAGGAACAGAAACAGCTGCCCTGGTCA CGGACCCAGCAGCAGCTAGAAA CAGCACTTAGT 7056
Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
QY 7057 AACCTACATCGAATGTGAACAGAGAGTCTCCAGACCCCTAGAAAATCTGTCACTAACCTG 7116
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 7117 GAGGAATCCCTAACCTCTTATCTGAAGTAGTCTCTACAGAAATAGAGAGGTTAGATTTA 7176
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgGlyLeuAspLeu 540
QY 7177 TTATTTCTAAAGAGAGGATTATGTGTAGCTTGAAGCCTTAAGAGAGAAATGCTTTTATGTG 7236
Db 541 LeuPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal 560
QY 7237 GATCATTACGGGCCCATCAGAGCTCCATGAA CAAGCTTAGAGAAAGGTTGGAGAACGCT 7296
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArg 580
QY 7297 CGAAGGAAAAGAAACTACTCAAGGCTGTTTGAGGATGGTTCAACAGGTCTCTTTGG 7356
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 7357 TTGGCTACCCCTACTTTCTGCTTTAAACAGGACCCCTTAATAGTCTCTCTCTTACTCACA 7416
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 620
QY 7417 GTTGGGCATGTAATTATTAACAAGTTAATGTCCTTCAATAGACAAGAAATAGTGCAATC 7476
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
QY 7477 CAGATCATGTACTTAGACAACAGTACCAA 7506
Db 641 GluIleMetValLeuArgGlnGlnTyrGln 650
```

RESULT 9

```
US-09-171-553B-4
; Sequence 4, Application US/09171553B
; Patent No. 6756227
; GENERAL INFORMATION:
; APPLICANT: GALBRAITH, DANIEL N.
; APPLICANT: LEES, GILLIAN M.
; APPLICANT: SMITH, KENNETH T.
; TITLE OF INVENTION: PORCINE RETROVIRUS
; FILE REFERENCE: CFV-5.01
; CURRENT APPLICATION NUMBER: US/09/171,553B
; CURRENT FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: PCT/GB97/01087
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: GB 9702668.6
; PRIOR FILING DATE: 1997-02-10
; PRIOR APPLICATION NUMBER: GB 9608164.1
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Porcine retrovirus
US-09-171-553B-4
```

```
Alignment Scores:
Pred. No.: 1,97e-217 Length: 524
Score: 2640.00 Matches: 495
Percent Similarity: 97.3% Conservative: 16
Best Local Similarity: 94.3% Mismatches: 13
Query Match: 18.0% Indels: 2
```

```
DB: 2 Gaps: 0
US-10-723-552-3 (1-8132) x US-09-171-553B-4 (1-524)
QY 585 ATGGGACAGACGGTGTAGCGACCCCTCTTAGTTTGACTCTCGACCATTTGGACTGAAGTTAAA 644
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20
QY 645 TCCAGGGCTCATTAATTTGTTCAGTTTCAGTTTAAAGAGGACCTTGGCAGACTTTCTGTGTC 704
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysGlyProTrpGlnThrPheCysAla 40
QY 705 TCTGAATGCCCGACCATTTGATGTCATCAGAGGGGACCTTTAATTTCTGAGATT 764
Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
QY 765 ATCTCGCTGTTTAAACAGATTATTTTTCAGACTGTGACCCGCTCTCATCTCCGATCAGGAG 824
Db 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
QY 825 CCTATATCTTACGTGGCAAGATTTTGGCAGAGGATCCTCGCCATGGGTTAAACCATGG 884
Db 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
QY 885 CTGAATAAGCCAAAGACCCAGGTCCCGAATTTCTGGCTCTTCGAGAGAAAAACAAACAC 944
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
QY 945 TCGGCTGAAAAAGTCAAGCCCTC-TCCTCATATCTACCCCGAGATTGAGGAGCCACCGGC 1003
Db 121 SerAlaGluLysValGluProSerSerTyrLeuProArgAspArgGlyAlaAlaAsp 140
QY 1004 TTGGCCGGAACCCCAATCTGTTCCCGACCCCTTATCTGGCACAGGTGCCCCGGGGG 1063
Db 141 LeuAlaGlyThrProThrCysSerProThrProLeuSerSerThrGlyCysGluGly 160
QY 1064 ACCCTTTGCCCTCTCGAGCTCCGGCGGTGGAGGACCTGCTGCGAGGACTCGAGCGCG 1123
Db 161 Thr-SerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerAr 180
QY 1124 GAGGGCGCCACCCCGGACGACAGATTCGCGACATTACCGCTGGCGACGTACGG 1183
Db 180 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTrpG 200
QY 1184 CCTCCCAACACGGGGGGCCAAATGCGACCCCTCCAGTAGTTGGCCCTTTTCTCTCAGA 1243
Db 200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerAlaAs 220
QY 1244 TCTCTATAATTGGAAAACTAACCATCCCCCTTTCTCGAGGATCCCAACGCTCACGGG 1303
Db 220 PheTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrG 240
QY 1304 GTTGTGGAGTCCCTTATGTTCTCACCAGCCTACTTGGGATGATGTCAACAGCTGCT 1363
Db 240 yLeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
QY 1364 GCGACACTCTTCAACAACCGAGGACGAGAGAAATCTATTAGAGCTAGAAAAATGT 1423
Db 260 uGlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVa 280
QY 1424 TCCTGGGGCGGACGGCGACCCACGCGTTTGAATAATGAGATTGACATGGGATTTCCCTT 1483
Db 280 lProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe 300
QY 1484 AACTCGCCCGGTTGGGACTTACAAACCGCTGAAGGTAGGAGAGCTTGAATAATCTATCG 1543
Db 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrAr 320
QY 1544 CAGGCTCTGGTGGCGGGTCTCCGGGGCGGCTCAAGACGGCCCACTAATTTGGCTAAGGT 1603
Db 320 gGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVa 340
QY 1604 AAGAGAAAGTGTGTCAGGGACCGGAATGAACCCCTCTGTTTTTCTTGAGAGGCTCTTGGGA 1663
```

```
Db 340 lArgGluValMetGlnGlyProAsnGluProSerValPheLeuGluArgLeuMetGl 360
Qy 1664 AGCTTCAGGCGGTACACCCCTTTTGATCCACCTCAGAGCCCCAAAAGCCTCAGTGGC 1723
Db 360 uAlaPheArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAl 380
Qy 1724 TTGGCCCTTTAGGACAGTCAGCCTTGGATATTAGAAAGCTTCAGAGACTGGAGG 1783
Db 380 aLeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
Qy 1784 GTTACAGGAGCTGAGTACGTGATCTAGTCAGGAGGACAGAAAGTATTACAAAG 1843
Db 400 yLeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArg 420
Qy 1844 GAGACAGAAAGAAAGGAAACAAAGAAAGAGAGAAAGAGAGAAAGGAGGAAAG 1903
Db 420 gGluThrGluGluGluLysGluGlnArgLysGluLysGluArgGluGluArg 440
Qy 1904 ACCTAATAACGGAAGAGAAATTTGACTAAGATCTTGGCTGCAAGTGTGAAAGGAA 1963
Db 440 gArgAspArgArgGlnGluLysLeuThrLysIleLeuAlaAlaValValGluGly 460
Qy 1964 AAGCAATACGAAAGAGAGAGATTTTAGGAAATTTAGTCCAGCCCTAGACAGTCAG 2023
Db 460 sSerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSer 480
Qy 2024 GAACCTGGGCAATAGAGCCCACTCGCAAGAGGACCAATGTGCATATTGTAAAGAAAGG 2083
Db 480 yAsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGlu 500
Qy 2084 ACCTGGGCAAGAACTGCCCCAGAGAGGAAACAAAGGACCAAGGATCCTAGCTTAGA 2143
Db 500 yHisTrpAlaArgAsnCysProLysGlyAsnLysGlyProLysValLeuAlaLeu 520
Qy 2144 AGAAGATAAGAT 2156
Db 520 uGluAspLysAsp 524

RESULT 10
US-08-929-967-7
; Sequence 7, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Rupert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-Length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 amino acids
```

```
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-929-967-7
Alignment Scores:
Pred. No.: 2,2e-205 Length: 665
Score: 2500.50 Matches: 460
Percent Similarity: 83.5% Conservative: 93
Best Local Similarity: 69.5% Mismatches: 106
Query Match: 17.1% Indels: 3
DB: 1 Gaps: 3

US-10-723-552-3 (1-8132) x US-08-929-967-7 (1-665)
Qy 2514 TTCAACCTCCCAATTAGATGACGAATATCGACTATATCTCTCCCTAGTAAAGCCTGATCAA 2573
Db 1 MetProLeuAsnIleGluAspGluHisArgLeuHisGluThrSerLysGluProAspVal 20
Qy 2574 AATATA--CAATTCTGGTTGGAAACAGTTTCCCAAGCCTGGCCAGAAACCCGAGGGATG 2630
Db 21 SerLeuGlySerThrTrpLeuSerAlaPheProGlnAlaTrpAlaGluThrGlyGlyMet 40
Qy 2631 GGTTCGGCAAGCAAGTTCCTCCCAAGTATTCAACTGAGGCCAGTCCACACCGATG 2690
Db 41 GlyLeuAlaValArgGlnAlaProLeuIleIleProLeuLysAlaThrSerThrProVal 60
Qy 2691 TCAGTCAGACAGTACCTCCCTTGAGTAAAGAGAGCTCAAGAGAGAAATTCGGCCGATGTC 2750
Db 61 SerIleLysGlnTyrProMetSerGlnGluAlaArgLeuGlyIleLysProHisIleGln 80
Qy 2751 AGATTAATCAACAGGGCATCTCTAGTTCCTGTCCTCAATCTCCCTGGAATACTCCCTGCTA 2810
Db 81 ArgLeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeuLeu 100
Qy 2811 CCGTTAGAAAGCCTGGGACTAATGACTATCGACAGTACAGACTTGAGAGAGGTCAAT 2870
Db 101 ProIleLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsn 120
Qy 2871 AAACGGGTGAGGATATACACCAACAGTCCCAAGACCTTATAACCTCTTGTGCTCTC 2930
Db 121 LysArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeu 140
Qy 2931 CCACCCCAACGGAGCTGGTATACAGTATTGGACTTAAAGGATGCCTTCTTCTCCCTGAGA 2990
Db 141 ProProSerHisGlnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArg 160
Qy 2991 TTACACCCCACTAGCCAAACCACTTTTGGCTTCGAATGAGAGATCCAGTACGGGAAGA 3050
Db 161 LeuHisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGluMetGlyIle 180
Qy 3051 ACGGGCAGCTCACCTGGACCCGACTGCCCCCAAGGGTTCAAGAACTCCCGACCATCTTT 3110
Db 181 SerGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPhe 200
Qy 3111 GACGAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCAAACCCCTCAGGTGACCCCTC 3170
Db 201 AspGluAlaLeuHisArgAspLeuAlaGlyPheArgIleGlnHisProAspLeuLeuLeu 220
Qy 3171 CTCCTAGCTAGTACCTGCTCTTGGGGGAGGCCAACAAACAGAGACTCTCTTAGAAGGC 3230
Db 221 LeuGlnTyrValAspAspLeuLeuAlaIleSerSerGluLeuAspCysGlnGlnGly 240
Qy 3231 ACGAAGGCACTACTGCTGAAATTGCTGACTAGGCTACAGAGCCTCTGCTAAGAGGCC 3290
Db 241 ThrArgAlaLeuLeuGlnThrLeuGlyAspLeuGlyTyrArgAlaSerAlaLysAla 260
Qy 3291 CAGATTTGCAGGAGAGAGTAAACATACTTGGGGTACAGTTTTCGGGACCGGAGCGATGG 3350
Db 261 GlnIleCysGlnLysGlnValLysTyrLeuGlyTyrLeuLeuLysGluGlyGlnArgTrp 280
Qy 3351 CTGACGGAGGACCGAAGAAACTGTAGTCAGATACCGGCCCCCAACCAACAGCAACAA 3410
Db 281 LeuThrGluAlaArgLysGluThrValMetGlyGlnProThrProLysThrProArgGln 300
```

```
QY 3411 ATGAGAGTGTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGTTCGGACC 3470
Db 301 LeuArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpLeuPheAlaGlu 320
QY 3471 TTAGCAGCCCACTCTACCGCTTAACCAAGAAAGGGAAATTCCTCGGCTCTCGAG 3530
Db 321 MetAlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAsp 340
QY 3531 CACAGAAGGATTTGATGTATCAAAAAGCCCTGTGTAGCGCACCTGTCTGTGGCCCTC 3590
Db 341 GlnGlnLysAlaTyrGlnGluLeuLysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeu 360
QY 3591 CTTGACGTAACTAAACCTTTACCTTTATGTGGAGCGTAAGGAGTAGCCCGGGA 3650
Db 361 ProAspLeuThrLysProPheGluLeuPheValAspGlnLysGlnGlyTyrAlaLysGly 380
QY 3651 GTTTTAAACCAACCCCTAGACCATGAGAGAAGACCTGTCCCTACCTGTCAAAAGAGCTC 3710
Db 381 ValLeuThrGlnLysLeuGlyProTrpArgProValAlaTyrLeuSerLysLysLeu 400
QY 3711 GATCTGTAGCCAGTGTGGCCCATATGCTGAAGGCTATCGCAGCTGTGGCCATAGT 3770
Db 401 AspProValAlaAlaGlyTrpProProCysLeuArgMetValAlaAlaLeuValLeu 420
QY 3771 GTCAAGGACGCTGACAAATGACTTTGGGACAGATATACTGTATAGCCCCCATGCA 3830
Db 421 ThrLysAspAlaGlyLysLeuThrMetGlnProLeuValLysLeuAlaProHisAla 440
QY 3831 TTGGAGAACATCGTTCGGCAGCCCCAGACCGATGATGACCAACGCCGCCCATGACCCAC 3890
Db 441 ValGluThrLeuValLysGlnProProAspArgTrpLeuSerAsnAlaArgMetThrHis 460
QY 3891 TATCAAGACCTGCTTCTC---ACAGAGAGGGTCACTGTCGCTCCACGCGCTCTCAAC 3947
Db 461 TyrGlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProValValAlaLeuAsn 480
QY 3948 CTTGCCACTCTTCTGCTGAAGAGACTGATGAACCGAGTACTCATGATTCCTCACTA 4007
Db 481 ProAlaThrLeuLeuPro---LeuProLysGluGlyLeuGlnHisAspCysLeuAspIle 499
QY 4008 TTGATTGAGGAGACTGGGGTCCGCAAGACCTTACAGACATACCGCTGACTGAGAGAAGTG 4067
Db 500 LeuAlaGluAlaHisGlyThrArgSerAspLeuThrAspGlnProLeuProAspAlaAsp 519
QY 4068 CTAACTGTTTCACTGACGAAAGCAGCTATGTGTGGAAAGGTAAAGGATGGCTGGGGCG 4127
Db 520 HisThrTrpTyrThrAspGlySerSerPheLeuGlnGlyGlnArgLysAlaGlyAla 539
QY 4128 GCGTGTGTGACGGGACCCGACGATCTGCGCCAGCAGCTGCGGAGGAAGCTTCAGCA 4187
Db 540 AlaValThrThrGluThrGluValIleTrpAlaArgAlaLeuProAlaGlyThrSerAla 559
QY 4188 CAAAAGGCTGAGTCACTGCGCTCACGCAAGCTTTGCGGCTGCGCGAAGGAAATCCATA 4247
Db 560 GlnArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysLysLeu 579
QY 4248 AACATTTATAGGACAGCAGGTATGCTTTGCGACTGCACAGTACATGGGGCCATCTAT 4307
Db 580 AsnValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyGluIleTyr 599
QY 4308 AAACAAGGGGTGTCTTACTCAGCAGGAGGGGAATTAAGACAAAGAGGAATTTCTA 4367
Db 600 ArgArgArgGlyLeuLeuThrSerGluGlyGluIleLysAsnLysGlyGluIleLeu 619
QY 4368 AGCTATTAGAAGCCGTACATTACCAAAAAGGCTAGCTATTATACATCTGCTGGACAT 4427
Db 620 AlaLeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHis 639
QY 4428 CAGAAAGCTAAGATCTCATATCCAGAGGAAACACAGATGCTGACCGGGTTGCCAAGCAG 4487
Db 640 GlnLysGlyAsnSerAlaGluAlaArgGlyAsnArgMetAlaAspGlnAlaAlaArgLys 659
```

```
QY 4488 GCAGCC 4493
Db 660 AlaAla 661

RESULT 11
US-09-111-085-4
; Sequence 4, Application US/09111085
; Patent No. 6100034
; GENERAL INFORMATION:
; APPLICANT: Stoye, Jonathan P
; APPLICANT: Weiss, Robin A
; TITLE OF INVENTION: Detection of retroviral subtypes based upon envelope
; FILE REFERENCE: 4238/75168
; CURRENT APPLICATION NUMBER: US/09/111,085
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: GB 9710154.7
; EARLIER FILING DATE: 1997-05-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Porcine retrovirus
US-09-111-085-4

Alignment Scores:
Pred. No.: 1,2e-198 Length: 657
Score: 2422.00 Matches: 457
Percent Similarity: 80.2% Conservative: 68
Best Local Similarity: 69.8% Mismatches: 96
Query Match: 16.5% Indels: 34
DB: 2 Gaps: 9

US-10-723-552-3 (1-8132) x US-09-111-085-4 (1-657)
QY 5620 ATGCATCCCAAGTAAACCGCGCCACCTCCCGATTCGGGGTGGAAAGCGAAAAAGACTG 5679
Db 1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyLeuProLysArgLeu 20
QY 5680 AAAATCCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCATAACTCTCTCAG 5739
Db 21 ArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrProGln 40
QY 5740 ACTAATGTATGGCATAGGAGACAGCTGAACCTCCATAAACCCTTATCTCACCCTGG 5799
Db 41 AlaSerSerLysArgLeuLeuAspSerSerAsnProHisArgProLeuSerLeuThrTrp 60
QY 5800 TTAATTTACTGACTCCGCGCACAGGTATTAAATATCAACAACACTCAAGGGGAGGCTCTTTA 5859
Db 61 LeuIleIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
QY 5860 GGAACCTGTGGCTGATCTATACGTTTGCCTCAGATCAGATTATTCCTAGTCTGACCTCA 5919
Db 81 GlyThrTrpTrpProGluLeuHisPheCysLeuArgLeuLeuAsnProAlaValLysSer 100
QY 5920 ----CCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCAACCAATAAT 5976
Db 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyThrGluLysGlu 120
QY 5977 GGAACAACATTGCGGAATCCAGAGATTTCTTTTGTAAACAATGGAACCTGTGAACCTCT 6036
Db 121 ---LysTyrCysGlyGlySerGlyGluSerPheCysArgArgTrpSerCysValThrSer 139
QY 6037 AATGATGGATTTGGAATGGCAACCTCTCAGCAGGATAGGTAAGTTTCTTATGTC 6096
Db 140 AsnAspGlyAspTrpLysTrpProIleSerLeuGlnAspArgValLysPheSerPheVal 159
QY 6097 AACACCTATACGACTCTGGCAATTAATTAACCTGACCTGGATAGTAACCTGGNAGCCCC 6156
Db 160 Asn-----SerGlyProGlyLysTyrLysValMetLysLeuTyrLysAspLys 175
QY 6157 AAGTGCTCTCTTTCAGACCTAGATTACCTAAAAATAAGTTTTCATCTGAGAAAGGAAACAA 6216
```


Db 61 LeuIleleAspProaspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
QY 5860 GGAACCTGGTGGCTGATATACGTTGGCTCAGATCAGTATATCTAGCTGACCTCA 5919
Db 81 GlyThrTrpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
QY 5920 ---CCCCCATATCCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCAACATAAT 5976
Db 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyThrGluLysGlu 120
QY 5977 GGAACCATGCGGAAATCCAGAGATTCTTTTGTAAACAATGGAACCTGTGAACCTCT 6036
Db 121 ---LysTyrCysGlyGlySerGlyGluSerPheCysArgArgTrpSerCysValThrSer 139
QY 6037 AATGATGATATTTGGAATAGCCCAACCTCTCAGCAGATAGGTAAAGTTTCTTATGTC 6096
Db 140 AsnAspGlyAspTrpLysTrpProLysSerLeuGlnAspArgValLysPheSerPheVal 159
QY 6097 AACACCTATACCACTCTGACAAATTAATTAACCTGACCTGGATTAGAACTGAAGCCCC 6156
Db 160 Asn-----SerGlyProGlyLysValMetLysLeuTyrLysAspLys 175
QY 6157 AAGTGCTCTCTTCAGACCTAGATTACTTAAATAAGTTTCTCAGAAAGGAAACAA 6216
Db 176 SerCysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGln 195
QY 6217 GAAATATCTTAAATGGGTAATGTTATCTTGGGGAATGTTATAT-----TATGGA 6270
Db 196 GluAsnIleGlnLysTrpLysAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGly 215
QY 6271 GGCTCGGTAACACACAGGCTCCATCTTAACCTTTCGCTCAAAATAAACACAG---CTG 6327
Db 216 GlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThr 232
QY 6328 GAGCCTCCAATGCTATAGACCAATAACCGTCTTGACGGTCAAAGACCCCAACCCAA 6387
Db 233 GluProProValAlaValGlyProAspLysValLeuAlaGluGlnGlyProProAlaLeu 252
QY 6388 GGAACA-----GGG-----GGA 6396
Db 253 GluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGln 272
QY 6397 CCATCCTCTAACATACTCTGGATCAGACCCCACTCAGTCT-----AACAGC----- 6444
Db 273 ProProSerAsnGlyThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 292
QY 6445 ---ACGACTAAATGGGGCAAACTTTTATGCTCATCCAGGAGCTTTTCAAGCTCTT 6501
Db 293 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 312
QY 6502 AACTCCACGACTCCAGAGGCTACTCTTCTTGTGGCTATGCTTAGCTTCGGGCCCACT 6561
Db 313 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 332
QY 6562 TACTATGAAGATTCGTAGAGGGAATTCATGTGACAAAGACATAGAGACCAA 6621
Db 333 TyrTyrGluGlyMetAlaLysGluGlyLysPheAsnValThrLysGluHisArgAsnGln 352
QY 6622 TGACATGGGGATCCCAAAATAAGCTTACCTTACTGAGGTTTCTGGAAAAGGCACCTGC 6681
Db 353 CysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCys 372
QY 6682 ATAGGAAGGTTCCCATCCCAACCAACACTTTGTAAACACACTGAAGCCTTTAATCAA 6741
Db 373 IleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 392
QY 6742 ACCTCTCAGATCAATATCGTACCTGGTATGACAGGTGGTGGCATGTATACTGGA 6801
Db 393 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpAlaCysAsnThrGly 412
QY 6802 TTAACCCCTTGTGTTTCCACCTTGGTTTAAACCAACTAAAGATTTTTCATTATGGTC 6861

Db 413 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMetVal 432
QY 6862 CAAATTTTCCCGAGTGTATTACTATCCGAAAAAGCAATCTTGTATGATATGACTAC 6921
Db 433 GlnIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAspTyr 452
QY 6922 AGAAATCATCGCAAAAGAGAGAACCATATCTCTGACACTTCTGTGATGCTCGGACTT 6981
Db 453 ArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu 472
QY 6982 GGAAGTGGCAGCAGCTAGAGAACAGAGACAGCTGCCCTGTGTACCGGACCAACAGACTA 7041
Db 473 GlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 492
QY 7042 GAAACAGGACTTAGTAACCTACATCGAATTTGTAACAGAGATCTCAAGCCCTAGAAAAA 7101
Db 493 GluLysGlyLeuGlyLeuHisAlaIleMetThrGluAspLeuArgAlaLeuGluGlu 512
QY 7102 TCTGTCAAGTAACCTGAGGAATCCCTAACTCTTATCTGAAGTAGTCTTACAGAATAGA 7161
Db 513 SerValSerAsnLeuGluGluSerLeuThrSerLeuValValLeuGlnAsnArg 532
QY 7162 AGAGGTTTAGATTATTTCTTAAAGAGGAGGATTTATGTGTAGCTTGAAGAGGAA 7221
Db 533 ArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluGlu 552
QY 7222 TGCTGTTTTTATGTGATCATTCAGGGGCCATCAGAGACTCCATGAAACAGCTTAGAGAA 7281
Db 553 CysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGlu 572
QY 7282 AGCTTCGAGAAGCGTCGAAGGAAAAAGAACTACTCAAGGGTGGTTTGAGGGATGGTTC 7341
Db 573 ArgLeuGluArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 592
QY 7342 AACAGTCTCTTTGGTGGCTACCTACTTTCTGTCTTAAACAGGACCTTAAATAGTCTC 7401
Db 593 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValLeu 612
QY 7402 CTCCTGTTACTCAGCTTGGCCCATGTATTTAAACAAGTTAATTCCTTCAATAGAGAA 7461
Db 613 LeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlu 632
QY 7462 CGAATAAGTCAGTCCAGATCATGTGTACTTTAGACAAACAGTACCAC 7506
Db 633 ArgValSerAlaValGlnIleMetValLeuArgGlnTyrGln 647

RESULT 13

US-09-171-553B-10
Sequence 10, Application US/09171553B
Patent No. 6756227
GENERAL INFORMATION:
APPLICANT: GALBRAITH, DANIEL N.
APPLICANT: HAWORTH, CHRISTINE
APPLICANT: LEE, GILLIAN M.
APPLICANT: SMITH, KENNETH T.
TITLE OF INVENTION: PORCINE RETROVIRUS
FILE REFERENCE: CFV-5.01
CURRENT APPLICATION NUMBER: US/09/171,553B
CURRENT FILING DATE: 1999-02-08
PRIOR APPLICATION NUMBER: PCT/GB97/01087
PRIOR FILING DATE: 1997-04-18
PRIOR APPLICATION NUMBER: GB 9702668.6
PRIOR FILING DATE: 1997-02-10
PRIOR APPLICATION NUMBER: GB 9608164.1
PRIOR FILING DATE: 1996-04-19
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 657
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Amino acid sequence

OTHER INFORMATION: of ENV region "Raji"
US-09-171-553B-10

Alignment Scores:

Pred. No.: 2 5e-196 Length: 657
Score: 2395.00 Matches: 451
Percent Similarity: 80.3% Conservative: 75
Best Local Similarity: 68.9% Mismatches: 95
Query Match: 16.4% Indels: 34
DB: 2 Gaps: 10

US-10-723-552-3 (1-8132) x US-09-171-553B-10 (1-657)

```
QY 5620 ATGCATCCACGTTAAACCGCGCCACCTCCCGATTCCGGGTGGAAGCGAAAGACTG 5679
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGluProLysArgLeu 20
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 5739
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 21 ArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleThrProGln 40
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5740 ACTAATGGTATGGCATAGAGACAGCTCGAATCCCATAAACCTTATCTCTCACCTGG 5799
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 41 AlaSerSerLysArgLeuIleAspSerSerAsnProHisArgProLeuSerProThrTrp 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5800 TTAATTAAGTATCCCGCACAGGTATTAATCAACCAACTCAAGGGAGGCTCCTTTA 5859
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 LeuIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5860 GGAACCTGGTGGCTGATCTATAGCTTGGCTCAGATCAGTATTCTCTGACTGACTCA 5919
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 81 GlyThrTrpTrpProGlnLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5920 ---CCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCAACAATAAT 5976
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 101 ThrProAsnLeuValArgSerTrpGlyPheCysProGlyThrGluLysGlu 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5977 GGAATAATTCGGAATCCAGAGATTTCTTTTAAACAATGGAACCTGTGAACCTCT 6036
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ---LysTrpCysGlySerGlyGluSerPheCysArgArgTrpSerCysValThrSer 139
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6037 AATGATGATATGGAATGGCAACTCTCAGCAGATAGGTAAGTTTCTTATGTC 6096
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 140 AsnAspGlyAspTrpLysTrpProIleSerLeuGlnAspArgValLysPheSerPheVal 159
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6097 AACACCTATACAGCTCTGACAAATTAATTAATCTGACCTGGATTAGAAGTGAACCCC 6156
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 160 Asn-----SerGlyProGlyLysTrpLysMetMetLysLeuTrpLysAspLysSer--- 176
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6157 AAGTGTCTCTCTTCAGACCTAGATTACCTAAAAATAAGTTTCACTGAGAAAGGAAACAA 6216
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 177 ---CysSerProSerAspLeuAspTrpLeuLysIleSerPheThrGluLysGlyGln 195
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6217 GAAATATCTCTAAATGGTAATGATGCTCTTGGGGATGGTATAT-----TATGGA 6270
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 196 GluAsnIleGlnLysTrpLysAsnGlyMetSerTrpGlyIleValPheTrpLysTrpGly 215
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6271 GGCTCGGGTAAACAACACGAGCTCCATTCTAATCTTCCGCTCAAAATAAACCCAG---CTG 6327
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 216 GlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThr 232
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6328 GAGCTCCCAATGCTATAGACCAATAATGCGTCTTGACGGGTCAAAAGACCCCAACCCAA 6387
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 233 GluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAlaLeu 252
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6388 GGACCA-----GGA 6396
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 253 GluProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGln 272
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6397 CCATCTCTTAACATACTTCTGATCAGACCCCACTGAGTCT-----AACAGC----- 6444
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 273 ProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 292
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 14

```
QY 6445 ---ACGACTAAATGGGGCAAAACATTTTGTAGCTCATCCAGGAGGCTTTTCAAGCTCTT 6501
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 293 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 312
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6502 AACTCCACGACTCCAGAGGCTACTCTTCTTGTGGCTATGCTTAGCTTCGGGCCCACT 6561
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 313 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 332
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6562 TACTATGAAGGAATGGCTAGACGAGGAAATTCATGTGTGACAAAGAACATAGAGACAA 6621
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 333 TyrTrpGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsnGln 352
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6622 TGCACATGGGGATCCCAATAAAGCTTACCTTACTGAGGTTCCTGGAAGGACCACTGC 6681
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 353 CysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCys 372
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6682 ATAGGAAGGTTCCCCCATCCCAACACACCTTTGTAAACCACTGNAGCCTTTAATCAA 6741
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 373 IleGlyLysAlaProProSerHisGlnHisLeuCysTrpSerThrValValTrpGluGln 392
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6742 ACCTCTGAGAGTCAATATCTGTACCTGGTTATGACAGGTGGTGGGCATGTAATACTGGA 6801
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 AlaSerGluAsnGlnTrpLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnThrGly 412
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6802 TTAACCCCTTGTGTTTCCACCTTGGTTTAAACCAACTAAAGATTTTTCATTTATGTC 6861
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 413 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspLeuCysValMetVal 432
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6862 CAATTTGTTCCCGAGTGTATCTATCCCGAAAAAGCAATCCTTGTATGAATATGACTAC 6921
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 433 GlnIleValProArgValTrpTrpHisProGluValValLeuAspGluTrpAspTrp 452
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6922 AGAATCATCGCAAAAGAGAACCCATATCTCTGACACTCTCTGTGTGTCTCGACTT 6981
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 453 ArgTrpAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu 472
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6982 GGAGTGGCAGCAGGTGTAGGAACAGGAACAGCTGCGCTGTGTCAACAGAGATCTCAAGCCCTAGAAAAA 7041
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 473 GlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 492
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7042 GAAACAGGACTTGTATTAACCTACATCGAATTTGTAACAGAGATCTCAAGCCCTAGAAAAA 7101
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 493 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGlu 512
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7102 TCTGTCAAGTAACCTGAGGAATCCCTAACCTCTTATCTGAAGTAGTCTTACAGAAATAGA 7161
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 513 SerValSerAsnLeuGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg 532
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7162 AGAGGTTAGATTATTATTCTTAAAGAGGAGGATTTGTGTAGCTTTGAAAGGAGGAA 7221
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 533 ArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluGlu 552
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7222 TCGTGTTTTATGTGGATCATTCAGGGGCATCAGAGATCCCATGAAACAGCTTAGAGAA 7281
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 553 CysCysPheTrpValAspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgLys 572
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7282 AGTGTGAGAAGCTCGAAGGAAAGAAACTACTCAAGGGTGTGTGAGGATGCTTC 7341
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 573 LysLeuLysArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 592
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7342 AACAGTCTCTTTTGGTGGCTACCTTACTTCTTCTTAAACAGGACCTTTAATAGTCCCTC 7401
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 593 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValLeu 612
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7402 CTCCTGTACTCAGAGTGGGCCATGTATTATTAACAAGTTAATTCCTTCATTAGAGAA 7461
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 613 LeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlu 632
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7462 CGAATAAGTGCAGTCCAGATCATGTACTTACACACACTACCAA 7506
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 633 ArgValSerAlaValGlnIleMetValLeuArgGlnTrpGln 647
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
US-09-376-781-3
; Sequence 3, Application US/09376781
; Patent No. 6261806
; GENERAL INFORMATION:
; APPLICANT: Banerjee, Papia T.
; APPLICANT: Patience, Clive
; APPLICANT: Anderson, Goran K.
; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
; Patent No. 6261806
; TITLE OF INVENTION: Use
; FILE REFERENCE: 61750-267
; CURRENT APPLICATION NUMBER: US/09/376,781
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/097,015
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; OTHER INFORMATION: sequence of a PERV env polypeptide previously
; OTHER INFORMATION: described
US-09-376-781-3

Alignment Scores:
Pred. No.: 1,01e-192 Length: 656
Score: 2353.00 Matches: 450
Percent Similarity: 79.8% Conservative: 74
Best Local Similarity: 68.5% Mismatches: 96
Query Match: 16.1% Indels: 37
DB: 2 Gaps: 10

US-10-723-552-3 (1-8132) x US-09-376-781-3 (1-656)
QY 5620 ATGCATCCCAAGTTAAACCGCGCCACCTCCGATTCGGGTGGAAAGCCGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgHisLeuProThrArgGlyGlyLeuProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTCAATAACCTCTCAG 5739
Db 21 ArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleThrProGln 40
QY 5740 ACTAATGGTATGGCATAGGAGACGCTGAATCCCATAAACCTTATCTCTCACCCTGG 5799
Db 41 AlaSerSerLysArgLeuIleAspSerSerAsnProHisArgProLeuSerLeuThrTrp 60
QY 5800 TTAATTACTACCTCCGCGCAGGTATTAATATCAACAACACTCAAGGGAGGCTCCTTTA 5859
Db 61 LeuIleIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
QY 5860 GGAACCTGGTGGCTGATCTATACGTTTGCTTCAGATTCAGTTATTCCTAGTCTGACCTCA 5919
Db 81 GlyThrTrpTrpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
QY 5920 ---CCGCCAGATACCTCCANGCTCAGGATTTTATGTTTGGCCAGGACCCACCAATAAT 5976
Db 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyThrGluLysGlu 120
QY 5977 GGAARACATTCGGAATCCAGAGATTTCTTTGTAAACAATGGAACCTGTGAACCTCT 6036
Db 121 ---LysTyrCysGlyGlySerGlyGluSerPheCysArgArgTrpSerCysValThrSer 139
QY 6037 AATGATGATTAATGGAAATCCCAACCTCTCAGCAGGATAGGATAGGTTTCTTATGTC 6096
Db 140 AsnAspGlyAspTrpLysTrpProIleSerLeuGlnAspArgValLysPheSerPheVal 159
QY 6097 AACACCTATACAGCTCTGACCAATTAATTAATCTACCTGGATAGAACTCGGAAGCCCC 6156
Db 160 Asn-----SerGlyProGlyLysTyrLysMetMetLysLeuTyrLysAspLysSer--- 176
```

```
QY 6157 AAGTGTCTCTCTTCAGACCTAGATTACCTAAATAAAGTTTCTACTGAGAAAGGAAACAA 6216
Db 177 ---CysSerProSerAspLeuAspTyrLeuLysIleSerPheThr--GluArgLysThrG 195
QY 6217 GAAATATCTCTAAATGGGTAATGTATGTCTTGGGGAAATGGTATATAT-----TATGGA 6270
Db 195 LysLysTyrSerLysValAspLysTrpTyrGluLeuGlyAsnSer-PheLeuLeuTyrGly 214
QY 6271 GGTCTCGGTAAACAACACGAGCTCCATTCTTAACCTATTTCGCTCAATAATAACACAG---CTG 6327
Db 215 GlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThr 231
QY 6328 GAGCTCTCAATGGCTATAGACCAATAACGCTCTTGACGGGTCAAAAGACCCCAACCAAC 6387
Db 232 GluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProAlaLeu 251
QY 6388 GGACCA-----GGA 6396
Db 252 GluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGln 271
QY 6397 CCATCTCTTAACATAACTTCTGGATCAGACCCCACTGAGTCT-----ACAGC----- 6444
Db 272 ProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291
QY 6445 ---ACGACTAAATCGGGCCAAACCTTTTAGCTCATCTCAGGGAGCTTTTCAAGCTCTT 6501
Db 292 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 311
QY 6502 AACTCCAGACTCCAGAGGCTACCTCTTCTTGTGTGGCTATGTCTTGGCGGCCACCT 6561
Db 312 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 331
QY 6562 TACTATGAGGATGCTAGAGGGAAATTCAAATGTGACAAAGAAACATAGACACCA 6621
Db 332 TyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsnGln 351
QY 6622 TGCAATCGGGATCCCAATAAAGTTACCTTACTGAGGTTTCTCGAAAGCCACCTGC 6681
Db 352 CysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCys 371
QY 6682 ATAGGAAAGGTTCCCTCCATCCCAACCAACCTTTGTAAACCACTGAAGCTTTTAATCAA 6741
Db 372 IleGlyLysAlaProProSerHisGlnHisLeuCysTyr-SerThrValValTyrGluGln 391
QY 6742 ACCTCTGAGAGTCAATATCTGTGCTGCTGTATGACAGGTGGGGCATCTAATACCTGGA 6801
Db 392 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGly 411
QY 6802 TTAACCCCTTGTGTTCCACCTTGGTTTAAACCAAACTAAAGATTTTTCATTATGGTC 6861
Db 412 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMetVal 431
QY 6862 CAAATTTGTTCCCGAGTGTTACTATCCGAAAAAGCAATCTCTGATGAATATGACTAC 6921
Db 432 GlnIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAspTyr 451
QY 6922 AGAAATCATCGCAAAAGAGAACCCATATCTCTGACACTTCTGTGATGCTCGGACTT 6981
Db 452 ArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu 471
QY 6982 GGAGTGGCAGCAGCTGTAGGAAACAGAAACAGCTGCGCTGTCACGGGACCCACAGCAGCTA 7041
Db 472 GlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
QY 7042 GAAACAGGACTTATAGTAACTTACATCGAAATGTTTAACAGAGATCTCCAAGCCCTAGAAAA 7101
Db 492 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGlu 511
QY 7102 TCTGTCACTAGTCACTCGAGGAACTCCCTAACTCTTATCTGAAGTACCTACAGATAGA 7161
Db 512 SerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg 531
QY 7162 AGAGGCTTAGATTATTATTCTTAAAGAGGAGGATTTATGTGTAGCTTTGAAAGGAGAA 7221
```

Db 532 ArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluGlu 551
QY 7222 TGCTGTTTATGTGATCAATCAGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAA 7281
Db 552 CysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAlaLysLeuArgLys 571
QY 7282 AGTTGAGAGAGCTCGAAGGAAAGAAAGAACTACTCAAGGGTGGTTTGAAGGATGCTC 7341
Db 572 LysLeuGluArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 591
QY 7342 AACAGTCTCTTGGTGGTACCTACTCTCTCTTACAGAGCCCTTAATAGTCTC 7401
Db 592 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeu 611
QY 7402 CTCTGTGTACTCACAGTGGCCCATGTATTATTAACAAGTAAATGGCTTCAATAGAGAA 7461
Db 612 LeuLeuLeuLeuThrValGlyProCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 631
QY 7462 CGAATAAGTCAGTCCAGATCATGGTACTTAGACAACAGTACCAG 7506
Db 632 ArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGln 646

RESULT 15

US-09-171-553B-6
; Sequence 6, Application US/09171553B
; Patent No. 6756227
; GENERAL INFORMATION:
; APPLICANT: GALBRAITH, DANIEL N.
; APPLICANT: HAWORTH, CHRISTINE
; APPLICANT: LSES, GILLIAN M.
; APPLICANT: SMITH, KENNETH T.
; TITLE OF INVENTION: PORCINE RETROVIRUS
; FILE REFERENCE: CFV-5.01
; CURRENT APPLICATION NUMBER: US/09/171,553B
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: PCT/GB97/01087
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: GB 9702668.6
; PRIOR FILING DATE: 1997-02-10
; PRIOR APPLICATION NUMBER: GB 9608164.1
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Porcine retrovirus
US-09-171-553B-6

Alignment Scores:
Pred. No.: 1,01e-192 Length: 656
Score: 2353.00 Matches: 450
Percent Similarity: 79.8% Conservative: 74
Best Local Similarity: 68.5% Mismatches: 96
Query Match: 16.1% Indels: 37
DB: 2 Gaps: 10

US-10-723-552-3 (1-8132) x US-09-171-553B-6 (1-656)

QY 5620 ATGCATCCAGGTTAAACCGGCCACCTCCGATTCGGGGTGAAGCCGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgHisLeuProThrArgGlyGlyProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGGTGTTCTTACTCTGTCAATACCTCTCAG 5739
Db 21 ArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrThrIleThrProGln 40
QY 5740 ACTAATGGTATGGCATAGGACAGCTGAATCCCATATAAACCCCTTATCTCTACCTCG 5799
Db 41 AlaSerSerLysArgLeuLeuLeuSerSerAsnProHisArgProLeuSerLeuThrTrp 60
QY 5800 TTAATTAAGTCCGGCAGAGTATTATATCAACAACACTCAAGGGGAGGCTCCTTTA 5859

Db 61 LeuIleIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
QY 5860 GGAACCTGGTGGCTGATCTATAGCTTGCCTCAGATCAGTTATTCTCTAGCTGACCTCA 5919
Db 81 GlyThrTrpProGluLeuHisPheCysLeuArgLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
QY 5920 ---CCCCAGATATCTCCATGCTCAGGATTTTATGTTTGCCTCAGGACCCACCAATAA 5976
Db 101 ThrProProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyThrGluLysGlu 120
QY 5977 GGAACAATTCGCGAATCCAGAGATTTCTTTTAAACAATGGAACCTGTGAACCTCT 6036
Db 121 ---LysTyrCysGlyGlySerGlyGluSerPheCysArgTrpSerCysValThrSer 139
QY 6037 AATGATGGATATTGGAATGCGCACTCTCAGCAGGATAGGCTAGTTTCTTATGTC 6096
Db 140 AsnAspGlyAspTrpLysTrpProIleSerLeuGlnAspArgValLysPheSerPheVal 159
QY 6097 AACACCTATACAGCTCTCGACAATTTAATTAATTAATTAATTAATTAATTAATTAATTA 6156
Db 160 Asn-----SerGlyProGlyLysTyrLysMetMetLysLeuTyrLysAspLysSer--- 176
QY 6157 AAGTGTCTCTCTCAGACCTAGATTAACCTAAATAAAGTTTCTACTGAGAAAGAAACAA 6216
Db 177 ---CysSerProSerAspLeuAspTyrLeuLysIleSerPheThr--GluArgLysThrG 195
QY 6217 GAAATATCTTAATAATGGTAAATGGTATGCTCTGGGGAATGGTATAT-----TATGGA 6270
Db 195 LysLysTyrSerLysValAspLysTrpTyrGluLeuGlyAsnSer--PheLeuLeuTyrGly 214
QY 6271 GGCTCGGGTAAACCAACACGAGCTCCATTCCTAATTAATTAATTAATTAATTAATTAATTA 6327
Db 215 GlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThr 231
QY 6328 GAGCTTCAATGGCTATAGGACCAATACGGCTTTCGCGGGTCAAGAGACCCCAACCCAA 6387
Db 232 GluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProAlaLeu 251
QY 6388 GGACCA-----GGA 6396
Db 252 GluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGln 271
QY 6397 CCATCTCTTAACATAACTTCTCGATCAGACCCCACTGAGTCT-----AACAGC----- 6444
Db 272 ProProSerAsnSerThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291
QY 6445 ---ACGACTAAATGGGCAAACTTTTACCTCATCCAGGAGCTTTTCAAGCTCTT 6501
Db 292 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 311
QY 6502 AACTCCAGACTCCAGAGGCTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6561
Db 312 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 331
QY 6562 TACTATGAAGGAATGGCTAGAAGAGGAAATTCATATGTGACAAAGAAACATAGACACCA 6621
Db 332 TyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsnGln 351
QY 6622 TCACATGGGATCCCAAAATTAAGCTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6681
Db 352 CysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCys 371
QY 6682 ATAGAAAGGTTCCCTCCACCAACACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6741
Db 372 IleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 391
QY 6742 ACCTCTGAGAGTCAATATCTGGTATGATGACAGTGGTGGCATGTAATACCTGCA 6801
Db 392 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGly 411
QY 6802 TTAACCCCTTGTGTTTCCACCTTGGTTTAAACCAAACTTAAAGATTTTTCATATGTC 6861

```
Db      412 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMetVal 431
Qy      6862 CAAATTGTTCCCGAGTGATTAATCTATCCGAAAGCAATCCTTGATGATATGACTAC 6921
      |||||
Db      432 GlnIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAspTyr 451
      |||||
Qy      6922 AGAAATCATCGACAAAGACAGACCCATATCTGTGACACTTGTGTGATGCTCGGACTT 6981
      ||| :|||
Db      452 ArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu 471
      |||
Qy      6982 GGAGTGGCAGCAGGTGAGGAAACAGGAAACAGCTGCCCTGTGTCACGGGACCACACAGCAGCTA 7041
      |||
Db      472 GlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 491
      |||
Qy      7042 GAAACAGGACTTAGTACCTACCTACGAAATTGTACAGAGATCTCCAGCCCTAGAAAAA 7101
      |||
Db      492 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGlu 511
      |||
Qy      7102 TCTGTCACTAACCTGGAGGAACTCCCTAACCTCTTATCTGAAGTAGTCTACAGAAATAGA 7161
      |||
Db      512 SerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg 531
      |||
Qy      7162 AGAGGGTTAGATTATTATTATTTTAAAGAGAGGAGATTATGTGTAGCCTTTGAAGGAGAA 7221
      |||
Db      532 ArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluGlu 551
      |||
Qy      7222 TGCTGTTTTTATGTGGATCATTCAGGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAA 7281
      |||
Db      552 CysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgLys 571
      |||
Qy      7282 AGGTTGGAGAGCGTCAAGGGAAGAACTACTCAAGGGTGGTTTGAGGGATGGTTC 7341
      |||
Db      572 LysLeuGluArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 591
      |||
Qy      7342 AACAGGTCTCTTTGGTTGGCTACCTACTTTCTGCTTTAAACAGGACCCCTTAATAGTCCTC 7401
      |||
Db      592 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeu 611
      |||
Qy      7402 CTCCTGTTACTCAGTTGGGCCAATGTATTATTAAACAAGTTAATTGCCCTTCATTAGAGAA 7461
      |||
Db      612 LeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlu 631
      |||
Qy      7462 CGAATAGTCAGTCCAGATCATGGTACTTAGACACAGTACCAA 7506
      |||
Db      632 ArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGln 646
```

Search completed: February 14, 2006, 16:15:06
Job time : 425.147 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:07:51 ; Search time 204.475 Seconds
(without alignments)
3323.427 Million cell updates/sec

Title: US-10-723-552-3
Perfect score: 14636
Sequence: 1 GCGTGTGTACGACTGTGGG.....CTGTTTGCATCAAAAAA 8132

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB/spool/US10723552/runat_14022006_125148_13205/app_query.fasta_1
-DB=Published_Applications_AA_Main -QFMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN 1.1.602 @runat_14022006_125148_13205 -NCPU=6 -ICPU=3
-NO_WRAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : Published Applications AA_Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3980.5	27.2	1199	4	US-10-677-558-2
2	3425	23.4	638	3	US-09-851-859A-6
3	3412	23.3	638	4	US-10-441-949-32
4	3412	23.3	638	4	US-10-441-949-34
5	3412	23.3	638	4	US-10-441-949-40
6	3412	23.3	638	4	US-10-441-949-42
7	3412	23.3	638	4	US-10-441-949-45
8	3318.5	22.7	678	4	US-10-441-949-36
9	3318.5	22.7	678	4	US-10-441-949-38
10	3147.5	21.5	653	4	US-10-029-656-2
11	3135.5	21.4	653	4	US-10-029-656-4

12	2879.5	19.7	661	4	US-10-441-949-43
13	2874.5	19.6	660	4	US-10-441-949-16
14	2871.5	19.6	660	4	US-10-441-949-8
15	2866.5	19.6	660	3	US-09-851-859A-5
16	2851	19.5	678	4	US-10-441-949-12
17	2849.5	19.5	660	4	US-10-441-949-10
18	2849.5	19.5	660	4	US-10-441-949-18
19	2835	19.4	678	4	US-10-441-949-14
20	2491	17.0	678	5	US-10-827-498-2
21	2487.5	17.0	716	3	US-09-845-157-2
22	2487.5	17.0	716	4	US-10-661-819-2
23	2485	16.8	678	4	US-10-441-949-26
24	2455	16.8	678	4	US-10-441-949-24
25	2432	16.6	657	4	US-10-441-949-22
26	2432	16.6	657	4	US-10-441-949-30
27	2432	16.6	658	4	US-10-441-949-44
28	2432	16.5	657	3	US-09-851-859A-4
29	2422	16.5	657	4	US-10-441-949-20
30	2422	16.5	657	4	US-10-441-949-28
31	2353	16.1	656	3	US-09-851-859A-3
32	2007	13.7	471	4	US-10-448-871A-31
33	1935	13.2	1438	5	US-10-450-763-42377
34	1867	12.8	471	4	US-10-448-871A-29
35	1862	12.7	471	4	US-10-448-871A-27
36	1850.5	12.6	470	4	US-10-448-871A-28
37	1846	12.6	471	5	US-10-448-871A-30
38	1821	12.4	1577	5	US-10-450-763-36834
39	1821	12.4	1577	5	US-10-450-763-57482
40	1784	12.2	473	4	US-10-448-871A-25
41	1768.5	12.1	472	4	US-10-448-871A-26
42	1572.5	10.7	1267	5	US-10-450-763-33781
43	1550.5	10.6	878	5	US-10-450-763-36843
44	1520	10.4	676	4	US-10-677-558-3
45	1515	10.4	817	5	US-10-450-763-40498

ALIGNMENTS

RESULT 1
US-10-677-558-2
; Sequence 2, Application US/10677558
; Publication No. US20040096972A1
; GENERAL INFORMATION:
; APPLICANT: AUDIT, Muriel
; APPLICANT: COSSET, Francois-Loic
; TITLE OF INVENTION: CHIMERIC PLASMID COMPRISING A REPLICATIVE RETROVIRAL GENOME AND
; FILE REFERENCE: 1759.135
; CURRENT APPLICATION NUMBER: US/10/677,558
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/FR02/03934
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: FR 0114976
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1199
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: product of pol gene of pAM plasmid
US-10-677-558-2

Alignment Scores:			
Pred. No.:	1.42e-256	Length:	1199
Score:	3980.50	Matches:	755
Percent Similarity:	76.2%	Conservative:	165
Best Local Similarity:	62.5%	Mismatches:	257
Query Match:	27.2%	Indels:	31
DB:	4	Gaps:	11

US-10-723-552-3 (1-8132) x US-10-677-558-2 (1-1199)

Db 720 ArgArgGlyLeuLeuThrSerGluGlyLysGluLeuLysAspGluLeuLeuAla 739
 QY 4371 CTATTAGAGCCGTACATTTTACCAAAAGGCTAGCTATTATATACACTGCTGCTGACATCAG 4430
 Db 740 LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHisGln 759
 QY 4431 AAGACTAAGATCTCATATCCAGAGAAACCATGCTGACCGGGTTGCCAAGACGCA 4490
 Db 760 LysGlyAsnSerAlaGluAlaArgGlyAsnArgMetAlaAspGlnAlaAlaArgGluVal 779
 QY 4491 GCC-----CAGGGTGTTAACTTCTGCTTATATAGAAATGCCAAAGCCCA 4538
 Db 780 AlaThrArgGluThrProGlyThrSerThrLeuLeuIleLysSerThrProTyrThr 799
 QY 4539 GAACCCAGACGACAGTACACCTTAGAAGACTGCAAGAGATAAAAGATAGACACAGTTC 4598
 Db 800 HisGluHisPheHisThrValThrAspThrLysAspLeuThrLysLeu----- 816
 QY 4599 TCTGAGACTCCGGAAGGAGCTCTCTATACCTCA-----GATGGG 4637
 Db 817 -----GlyAlaThrTyrAspSerAlaLysLysTyrTrpValTyrGlnGly 831
 QY 4638 AAGGAATCTGCCCCACAAAGAGGTTAGATATGTCCACAGATACATCGTCTAAC 4697
 Db 832 LysProValMetProAspGlnPheThrPheGluLeuLeuAspPheLeuHisGlnLeuThr 851
 QY 4698 CACCTAGGAACCTAAACACCTGCAGCAGCTGGTGCAG-----ACATCCCTTATCATGTT 4751
 Db 852 HisLeuSerPheSerLysThrLysAlaLeuLeuGluArgSerProSerProTyrTyrMet 871
 QY 4752 CTGAGGCTACAGGAGTGTGACTCGGTGGTCAAACTGTGTGCTGCCAGCTGGTT 4811
 Db 872 LeuAsnArgAspArgThrLeuLysAsnIleThrGluThrCysLysAlaCysAlaGlnVal 891
 QY 4812 ATGTCTAATCTCCAGATGCTCCAGGAGAGACTAAGCGGAGACCCAGCGCT 4871
 Db 892 AsnAlaSerLysSerAlaValLysGlnGlyThrArgValArgGlyHisArgProGlyThr 911
 QY 4872 CACTGGGAAGTGCATCTACTAGGTAAAGCCGGCTAAATACGAAACAAATACCTATTG 4931
 Db 912 HisTrpGluLeuAspPheThrGluValLysProGlyLeuTyrGlyTyrLysThrLeuLeu 931
 QY 4932 GTTTTGTAGACACCTTTTCCAGATGGGTAGAGCTTATCTCTAAGAAAGACTTCA 4991
 Db 932 ValPheValAspThrPheSerGlyTrpIleGluAlaPheProThrLysLysGluThrAla 951
 QY 4992 ACGGTGGTGGCTAAATACTGGAAGAAATTTTCCAGATTTGGAATACCTAAGTGA 5051
 Db 952 LysValValThrLysLysLeuLeuGluGluIlePheProArgPheGlyMetProGlnVal 971
 QY 5052 ATAGGTCAGACATCGTCCAGCTTTTGTGTCAGGTAAAGTCCAGGACTGGCCAGATA 5111
 Db 972 LeuGlyThrAspAsnGlyProAlaPheValSerLysValSerGlnThrValAlaAspLeu 991
 QY 5112 TTGGGATTCATTTGGAACCTGATGTCATACAGACCCCAAGCTCAGGACAGGTAGAG 5171
 Db 992 LeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGlu 1011
 QY 5172 AGGTAAGTAAAGACCACTTAAAGACCCCTTACTAATTCACCGGAGACTGGCTTAAT 5231
 Db 1012 ArgMetAsnArgThrIleLysGluThrLeuThrLysLeuAlaThrGlySerArg 1031
 QY 5232 GATTGGATAGCTCTCTGCTGCTTTTGTGCTTTTGTAGGTTAGGAACACCCCTGACAGTTT 5291
 Db 1032 AspTrpValLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThrProGlyProHis 1051
 QY 5292 GGGCTGACCCCTATGAATTAATCTACGGGGGACCCCTTGTAGTAAATTTGCTTCT 5351
 Db 1052 GlyLeuThrProTyrGluLeuLeuTyrGlyAlaProProProLeuValAsnPheProAsp 1071
 QY 5352 GTACATAGTGTGAGCTGCTGCTTCCAGCCCTTTGCTTCTAGGCTCAAGGACTTGAG 5411
 Db 1072 ProAspMetThrArgValThrAsnSerProSerLeuGlnAlaHisLeuGlnAlaLeuTyr 1091

QY 5412 TGGGTGAGACAACGACGCTGGAGCAACTCCGGAGGCGCTTACTCAGGAGGAGACTTG 5471
 Db 1092 LeuValGlnHisGluValTrpArgProLeuAlaAlaTyrGlnGluGlnLeuAspArg 1111
 QY 5472 CAG---ATCCACATCGTTTCCAGTGGGAGATTCACTTACGTTAGAGCCACCGTGA 5528
 Db 1112 ProValValProHisProTyrArgValGlyAspThrValTrpValArgHisGlnThr 1131
 QY 5529 GAAACCTCGAGACTCGGTGGAAGGCGCTTATCTGCTACTTTTGTGACACCAACCGCT 5588
 Db 1132 LysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThrThrProThrAla 1151
 QY 5589 GTGAAGTCAAGGAAATCTCCACCTGGATCCATGCATCCACGTTAAACCGGCGCCACCT 5648
 Db 1152 LeuLysValAspGlyIleAlaAlaTyrIleHisAlaAlaHisValLysAlaAlaAspThr 1171
 QY 5649 -----CCGATTCGGG-----TGAAGCGGAAAGACTGCAAAATCCCTTAAG 5693
 Db 1172 GluSerGlyProSerSerGlyArgThrTrpArgValGlnArgSerGlnAsnProLeuLys 1191
 QY 5694 CTTGCGCTCATCGCGTGGTTCCT 5717
 Db 1192 IleArgLeuThrArgGlySerPro 1199

RESULT 2

US-09-851-859A-6
 ; Sequence 6, Application US/09851859A
 ; Patent No. US20020065407A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Banerjee, Papi T.
 ; APPLICANT: Patience, Clive
 ; APPLICANT: Andersson, Goxan K.
 ; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
 ; TITLE OF INVENTION: Use
 ; FILE REFERENCE: 61750-321
 ; CURRENT APPLICATION NUMBER: US/09/851,859A
 ; PRIOR FILING DATE: 1999-08-18
 ; PRIOR APPLICATION NUMBER: 60/097,015
 ; PRIOR FILING DATE: 1998-08-18
 ; PRIOR APPLICATION NUMBER: US 09/376781
 ; PRIOR FILING DATE: 2000-08-18
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 638
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:PERV-C
 ; OTHER INFORMATION: polypeptide sequence taken from GenBank Accession
 ; OTHER INFORMATION: No. US20020065407A1 AF038600 for comparison.
 US-09-851-859A-6

Alignment Scores:
 Pred. No.: 1,37e-219 Length: 638
 Score: 3425.00 Matches: 637
 Percent Similarity: 100.0% Conservatives: 1
 Best Local Similarity: 99.8% Mismatches: 0
 Query Match: 23.4% Indels: 0
 DB: 3 Gaps: 0

US-10-723-552-3 (1-8132) x US-09-851-859A-6 (1-638)

QY 5620 ATGCATCCACGCTTAAACCGGCGCCACCTCCCGATTCGGGTGGAAAGCCGAAAGACTG 5679
 Db 1 MetHisProThrLeuAsnArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
 QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGCTCCATCGCTGCTTCTTACTCTGTCAATAACCTCTCAG 5739
 Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
 QY 5740 ACTAATGGTATGCGCATAGGACAGCCTGAACTCCCAATAAACCTTATCTCTCACCTGG 5799

Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTip 60
Qy 5800 TTAATTAAGTACCTCCGACAGAGTATTAATATACAGAACTACAGGGAGGCTCTTTA 5859
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
Qy 5860 GGAACCTGGTGGCTGATCATAGTTTGGCTCAGATCAGTATTCCTAGTCTGACCTCA 5919
Db 81 GlyThrTrpProAspLeuThrValCysLeuArgSerValIleProSerLeuThrSer 100
Qy 5920 CCCCCAGATATCTCTCCATGCTCAGGATTTTATGTTGCCAGGACACCAAAATAATGGA 5979
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
Qy 5980 AAACATTGCGAAATCCAGAGATTTCTTTTGTAAACATGGAAGTGTAACTCTAAT 6039
Db 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
Qy 6040 GATGGATATTGGAATGGCCAACTCTCAGCAGATAGGTAAAGTTTCTTATGTCAAC 6099
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn 160
Qy 6100 ACCTATACCACTCTGCAAAATTAATTAACCTGACCTGGATTAGAATGGAAGCCCCAAG 6159
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
Qy 6160 TGCTCTCTTCAGACCTAGATTACCTAAATAATAGTTTCTCAGAAAGGAAACAAGAA 6219
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
Qy 6220 AATATCTTAAATGGGTAAATGTTATGTTCTTGGGAATGTTATATTAGGAGCTGGGT 6279
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrGlyGlySerGly 220
Qy 6280 AAACAACAGGCTCCATTTCAACTATTTCGCTCAAAATAAACAGCTGGAGCTCCAAATG 6339
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
Qy 6340 GCTATAGGACCAATACGGTCTTGACGGGTCAAGACCCCAACCCCAAGACACGACCA 6399
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
Qy 6400 TCCTCTAACATACTCTGGATCAGACCCCACTCAGTCTAACAGCAGCAGCTAAATGGGG 6459
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerSerThrThrLysMetGly 280
Qy 6460 GCAAACTTTTACCTCATCCAGGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAG 6519
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
Qy 6520 GCTACCTCTTCTTGGCTATGCTTAGCTTCGGGCCACCTTACTACTATGAAGNATGGCT 6579
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMetAla 320
Qy 6580 AGAAGAGGGAATTCATATGTGACAAAGAACATAGACCAATGCACATGGGATCCAA 6639
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlyLysGln 340
Qy 6640 AATAAGCTTACCTTACTAGGTTTCTCGAAAAGGCACCTGCATAGGAAAGGTTTCCCCCA 6699
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
Qy 6700 TCCCACCAACACCTTTGTACCACTGAAGCCTTTAATCAACCTCTGAGAGTCAATAT 6759
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
Qy 6760 CTGGTACCTGTATGACGGTGGGCGCATGTAATACTGGATTAAACCCCTTGTTTCC 6819
Db 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
Qy 6820 ACCTTGGTTTTTAAACCAAACTAAAGATTTTGGATTATGGTCCAAATTTGTTCCCGAGTG 6879

Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
Qy 6880 TATTACTATCCCAAAAGCAATCCTTGTATGATATGACTACAGAAATCATCCACAAAAG 6939
Db 421 TyrTyrTyrProGlnLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
Qy 6940 AGAAGAACCCATATCTCTGACACTTGTGTGATCTCGGACTTCGGAGTGGCAGCAGGTGTA 6999
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
Qy 7000 GGAACAGGAACAGCTCCCTGGTCACGGGACCAACAGAGCTAGAAACAGAGCTTAGTAAC 7059
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
Qy 7060 CTACATCGAATGTCTAACAGAGATCTCCAGCCCTAGAAAACTCTGCTACGTAACTCGGAG 7119
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
Qy 7120 GAATCCCTAACCTCTTATCTGAAGTAGTCTCTACAGAAAGAGAGGTTAGATTATTA 7179
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
Qy 7180 TTTCTAAAAGAGAGAGATATATGTGTAGCTTGAAGAGGAATGCTGTTTTTATGTGGAT 7239
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
Qy 7240 CATTCAAGGGCCATCAGAGACTCCATGMAACAAGCTTAGAGAAAGTTGGAGAACGTCGA 7299
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArg 560
Qy 7300 AGGCAAAAGGAAACTACTCAAGGGTGGTTTGAGGGATGGTTCAACAGGCTCTCTTTGGTTG 7359
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
Qy 7360 GCTACCTACTTTCTGCTTTAACAGACCTTAAATAGTCTCTCTCTCTGTTACTCACAGTT 7419
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
Qy 7420 GGGCCATGATTATTAAACAGTTAATTGCCTTCACTTAGAGAACGAATAAGTCAGTCCAG 7479
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
Qy 7480 ATCATGGTACTTACAGAACAGTACCAGGCCCCCTCTAGCAGGGAAGCTGCCCGC 7533
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 3
US-10-441-949-32
; Sequence 32, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-32
Alignment Scores:

Pred. No.: 1.01e-218 Length: 638
 Score: 3412.00 Matches: 635
 Percent Similarity: 99.7% Conservative: 1
 Best Local Similarity: 99.5% Mismatches: 2
 Query Match: 23.3% Indels: 0
 DB: 4 Gaps: 0

US-10-723-552-3 (1-8132) x US-10-441-949-32 (1-638)

QY 5620 ATGCATCCCAAGTTAAACCGCGCCACCTCCGATTCGGGGTGAAGCGAAAGACTG 5679
 DB 1 MetHisProThrLeuSerArgHisLeuProIleArgGlyGlyProLysArgLeu 20
 QY 5680 AAAATCCCTTAAGCTTCGCTCATCGCTGCTTCTTACTCTGCTCAATACCTCTCAG 5739
 DB 21 LysIleProLeuSerPheAlaSerIleAlaIleThrPheLeuThrLeuSerIleThrSerGln 40
 QY 5740 ACTAATGGTATGGCATAGGACAGCCTCAATCCCATAAACCTTATCTCTCACTGG 5799
 DB 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
 QY 5800 TTAATTAAGTCTCCGCGCACAGGTATTATATCAACAACACTCAAGGGGAGGCTCTTTA 5859
 DB 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
 QY 5860 GGAACCTGGGCTGATCTATACGTTTGCTCAGATCAGTTATCTAGTCTGACCTCA 5919
 DB 81 GlyThrTrpProAspLeuThrValCysLeuArgSerValIleProSerLeuThrSer 100
 QY 5920 CCCCAGATATCTCCATGCTCAGGATTTTATGTTGGCCAGGACCAACCAATATGGA 5979
 DB 101 ProProAspIleLeuHisAlaHisGlyPheThrValCysProGlyProAsnAsnGly 120
 QY 5980 AAACATTGCGGAATCCAGAGATTTCTTTGTAAACAATGGAACCTGTAACTCTAAT 6039
 DB 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
 QY 6040 GATGGATATGGAAATGGCAACCTCTCAGCAGGATAGGTAAAGTTTCTTATGCAAC 6099
 DB 141 AspGlyTy-TrpLysTrpProThrSerGlnAspArgValSerPheSerTyValAsn 160
 QY 6100 ACCTATACCACTCTCGACATTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6159
 DB 161 ThrTyThrSerSerGlyGlnPheAsnTyLeuThrTrpIleArgThrGlySerProLys 180
 QY 6160 TGCTCTCTCAGACCTAGATTACCTAAATAAAGTTTCACTGAGAAAGCAACAAGAA 6219
 DB 181 CysSerProSerAspLeuAspTyLeuLysIleSerPheThrGlnLysGlyLysGlnGlu 200
 QY 6220 AATATCTAAATGGGTAAATGGTATGCTCTGGGGAATGGTATATTATGGAGCTCGGT 6279
 DB 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyTyTyGlySerGly 220
 QY 6280 AAACACCAAGCTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6339
 DB 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProMet 240
 QY 6340 GCTATAGGACCAATACGCTTGACGGGTCAAGACCCCAACCAAGGACGAGACCA 6399
 DB 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
 QY 6400 TCCTCTAACATACTTCTGGATCAGACCCCACTGAGTCTAAACAGCACGACTAAATGGGG 6459
 DB 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
 QY 6460 GCMAAATTTTACCTCATCCGGGAGCTTTTCAAGCTCTTAACTCCACGACTCCAGAG 6519
 DB 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
 QY 6520 GCTACCTCTTCTGTGGCTATGCTTAGCTTCGGGCCCACTTACTATCAAGGAATGGCT 6579
 DB 301 AlaThrSerSerCysThrLeuCysLeuAlaLeuGlyProProTyTyTyGlyMetAla 320

QY 6580 AGAAGAGGGAATTTCAATGTGACAAAGAACATAGAGACCAATGCAATGGGATCCCAA 6639
 DB 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
 QY 6640 AATAAGCTTACCTTACTGAGTTTCTGGAAGAGGACCTGCATAGGAAGGTTCCCCCA 6699
 DB 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
 QY 6700 TCCCAACCAACACCTTTGTAACCACTGAAGCCCTTAAATCAAACTCTGAGAGTCAATAT 6759
 DB 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTy 380
 QY 6760 CTGGTACCTGGTTATGACAGGTGGTGGCATGTAATACATGATTAACCCCTTGTGTTTC 6819
 DB 381 LeuValProGlyTyTrpAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
 QY 6820 ACCTGGTTTAAACCAAACTAAAGATTTTGCATTTATGGTCCAAATTTGTTCCCGAGTG 6879
 DB 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
 QY 6880 TATTACTATCCCGAAAAAGCAATCTTGATGATATGACTACAGAAATCATCGACAAAG 6939
 DB 421 TyTyTyTyProGlnLysAlaIleLeuAspGluTyAspTyArgAsnHisArgGlnLys 440
 QY 6940 AGAAGAACCATATCTCTGACACTTGTCTGCTCGGACTTGGAGTGGCAGCAGGTGA 6999
 DB 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaGlyVal 460
 QY 7000 GGAACAGGACAGCTCCCTGTCAGGACCCACACAGCAGCTAGAAAACAGACTTAGTAAC 7059
 DB 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
 QY 7060 CTACATTCGAATTTGTAACAGAGATCTCCAAGCCCTAGAAAAATCTCTAGTAACTGGAG 7119
 DB 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
 QY 7120 GAATCCCTAACCTCTTATCTGAAGTAGTCTTACAGATAGAGAGGGTTAGATTATTA 7179
 DB 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeu 520
 QY 7180 TTTCTAAAAGAGGAGGATTTATGTAGCTTGAAGGAGGAATGCTGTTTATGTTGAT 7239
 DB 521 PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyValAsp 540
 QY 7240 CATTCAGGGCCCATCAGACACTCCATGAACAGCTTAGAGAAAGGTTGAGAGAGCTCGA 7299
 DB 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuLysArgArg 560
 QY 7300 AGGGAAGGAAACTACTCAAGGGTGGTTTGAAGGATGGTTCAACAGGCTCTCTTGGTTG 7359
 DB 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
 QY 7360 GCTACCTACTTTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTCTGTTACTCAGTT 7419
 DB 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
 QY 7420 GGGCCATGATTATTAAACAAGTTAATTCCTTTCATTAGAGAACGATTAATGTCAGTCCAG 7479
 DB 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
 QY 7480 ATCATGGTACTTAGACCAACAGTACCAGCCCGCTCTAGCAGGGAAGCTGGCCGC 7533
 DB 621 IleMetValLeuArgGlnGlnTyGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 4

US-10-441-949-34
 ; Sequence 34, Application US/10441949
 ; Publication No. US20040116684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rodrigo, Allen
 ; APPLICANT: Ross, Howard A.
 ; APPLICANT: Mullins, James I.
 ; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES


```
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
Qy 7180 TTCTCTAAAGAGAGGATTATGTGTAGCCTTGAGGAGGAATGCTGTTTTATATGGAT 7239
Db 521 PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
Qy 7240 CATTGAGGGCCATCAGACTCCATCAACAAGCTTAGAGAAAGTTGGAGAGCGTCGA 7299
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
Qy 7300 AGGGAAGAGAACTACTCAAGGGTGTGTTGAGGATGGTTCAACAGGTCCTCTTGGTTG 7359
Db 561 ArgGluLysGluThrThrGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrLeu 580
Qy 7360 GCTACCTCTCTTCTGCTTAAACAGGACCTTAAATAGTCTCTCTGTTTACTACAGTT 7419
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
Qy 7420 GGGCCATGTATTATTACRAGTTAATTCGCTTCATTAGACGACGAATAAGTCAGTCCAG 7479
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
Qy 7480 ATCATGCTACTTACACAAAGTACCAAAGCCGCTCTAGCAGGGAAGCTGGCCGC 7533
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 6
US-10-441-949-42
; Sequence 42, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-42

Alignment Scores:
Pred. No.: 1,01e-218 Length: 638
Score: 3412.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 23.3% Indels: 0
DB: Gaps: 0

US-10-723-552-3 (1-8132) x US-10-441-949-42 (1-638)
Qy 5620 ATGCATCCACGTTAAACCGCGCCACCTCCGATTCGGGTGGAAGCCGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
Qy 5680 AAAATCCCTTAGCTTCGCTCCATCGCTGGTTCTTACTCTGTCGAATAAAGCTCTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaIlePheLeuThrLeuSerIleThrSerGln 40
Qy 5740 ACTAATGGTATGCCATAGGAGACAGCGTCAACTCCCATAAACCCCTTATCTCACCTGG 5799
```



```
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
Qy 6640 AATAAGCTTACCTTACTGAGGTTTCTGGAAGGACCTGCATAGAGAAAGTTGCCCCA 6699
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
Qy 6700 TCCACCAACACCTTTGTACCACTGAGCCCTTTAATCAACCTCTGAGTCAATAT 6759
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
Qy 6760 CTGCTACTCTGTTATGACAGCTGGTGGCATGTAACTAGTAACTCCCTTGTTTCC 6819
Db 381 LeuValProGlyTyrAspAsgTrpTpAlaCysAsnThrGlyLeuThrProCysValSer 400
Qy 6820 ACCTGGTTTTTAAACCAAACTAAAGATTTTGCATTATGCTCCAAATTTGTTCCCGAGTG 6879
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
Qy 6880 TATTACTATCCGAAAAGCAATCCTTGATGATATATGACTACAGAAATCATCGACAAAAG 6939
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
Qy 6940 AGAAGACCCATATCTCTGACACTTGCTGTGATGCTCGACTTGGAGTGGCAGCAGGTGTA 6999
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGlyVal 460
Qy 7000 GGAACAGGAACAGCTGCCGTGTCACGGGACCAAGAGCTAGAAAAGCAAGCACTTAGTAAC 7059
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
Qy 7060 CTACATCGAATTGTACAGAGATCTCAAGCCCTAGAAAATCTGTCACTAACCTGGAG 7119
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
Qy 7120 GAATCCCTTAACCTCTTATCTGAAGTAGTCTCAGAAATAGAGAGGTTAGATTATTA 7179
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
Qy 7180 TTTCTAAAAGAGGAGATTATGTGTAGCTTGAAGAGGAATGCTGTTTTTATGTGAT 7239
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
Qy 7240 CATTGAGGGCCATCAGAGACTCCATGAAACAAGCTTAGAGAAAGTTGGAGAGCGCTGA 7299
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
Qy 7300 AGGAAAAGGAACTACTCAAGGTTGTTTGGAGGATGGTTCAACAGGTCTCTTTGGTTG 7359
Db 561 ArgGluLysGluThrThrGlnGlyTyrPheAsnArgSerProTrpLeu 580
Qy 7360 GCTACCCCTACTTTCTGTTTAAACAGGACCTTAAATAGTCTCTCTGTTTACTACAGTT 7419
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrVal 600
Qy 7420 GGCCCATGTTATTAAACAAGTTAATTGCTTCAATTAGACAGCAAGATAAGTCAGTCCAG 7479
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
Qy 7480 ATCATGTTACTTACACAAAGTACCAAGCCGCTAGCAGGGAAGCTGCCCGC 7533
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638
```

RESULT 8

```
US-10-441-949-36
; Sequence 36, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
```

```
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-36
```

```
Alignment Scores:
Pred. No.: 1,8e-212 Length: 678
Score: 3318.50 Matches: 636
Percent Similarity: 93.8% Conservative: 1
Best Local Similarity: 93.7% Mismatches: 3
Query Match: 22.7% Indels: 39
DB: 4 Gaps: 8
```

US-10-723-552-3 (1-8132) x US-10-441-949-36 (1-678)

```
Qy 5620 ATGCATCCCACTGTTAAACCGCGCCACCTCCCGATTCCGGGTGGAAGCCGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu 20
Qy 5680 ARAATCCCTTAAGCTTCGCTCCATCGCTGCTTCTTACTCTGTCATAATAACTCTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
Qy 5740 ACTAATCGGTATGCGCATAGAGACAGCCTGAACTCCCATAAACCTTATCTCTCACCTGG 5799
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
Qy 5800 TTAATTACTGACGCCGACAGGTATTATATACAACTCAAGGGAGGCTCTTTTA 5859
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
Qy 5860 GGAACCTGGTGGCTCATCTATACGTTTGCTCAGATCAGTATTCTTCTAGTCTG----- 5913
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuAsnAsp 100
Qy 5914 ---ACCTCACCCCCAGATATCTCCATGTCTCAGGATTTTATTTGTTGCCCAGGACCA 5970
Db 101 GlnThrSerProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProPro 120
Qy 5971 AATAATGGAATAATTCGCGAAATCCAGAGATTCTTTTGTAAACAACTGGAAGTGTGTA 6030
Db 121 AsnAsnGlyLysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysVal 140
Qy 6031 ACCTCTAATGATCGATATTGGAATGCGCAACCTCTCAGCAGGATAGGGTAAGTTTCT 6090
Db 141 ThrSerAsnAspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSer 160
Qy 6091 TATGTCAACACCTTATACCACTCTGGACCTGGAATTTAATTATAC-----CTGACC 6135
Db 161 TyrValAsnThrTyrThrSerSerGlyGlnPheAsnTyrGlyHisGlyAsgTrpLeuThr 180
Qy 6136 TGG-----ATTAGAACTGGAAGCCCAAGTGTCTCTCTTCAGAC 6174
Db 181 TrpGlnGlnArgValGlnLysAspIleArgThrGlySerProLysCysSerProSerAsp 200
Qy 6175 CTAGATTACCTAAATAAGTTTCACTGAGAAAAGGAAACAAAGAAATATCTCTAAAATGG 6234
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleLeuLysTrp 220
Qy 6235 GTAAATGGTATGTCTTTGGGGAATGGTATATTATTTGAGGAGCTCGGGTAAACACAGGCTCC 6294
Db 221 ValAsnGlyMetSerTrpGlyMetValTyrTyrGlySerGlyLysGlnProGlySer 240
```

QY 6295 ATTCTAACTATTGCTCCCTCAAAATAAAC---CAGCTGGAGCCTCAATGGCTATAGGACCA 6351
Db 241 ILeuThrIleArgLeuLysIleAsnThrGlnLeuGluProProMetAlaIleGlyPro 260
QY 6352 AATACGGCTCTGACGGCTCAAGACCCCAACCCCAAGACCA-----6393
Db 261 AsnThrValLeuThrGlyGlnArgProProThrGlnGlyProProHisAsnLeuProVal 280
QY 6394 -----GGACCATCTCT-----AACATAACTTCTGGA 6420
Db 281 ProGlnGlyProSerProAsnProAspIleThrGlnSerAspTyrAsnIleThrSerGly 300
QY 6421 TCAGACCCCACT-----GAGTCTAACAGCAGCACTAAATGGGGCAAAA 6465
Db 301 SerAspProThrAsnThrProArgAsnGluSerAsnSerThrThrLysMetGlyAlaLys 320
QY 6466 CTTTATTAGCCTCATCCAGGAGCTTTCAAGCTCTTAAGTCCAGGCTCCAGAGGCTACC 6525
Db 321 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThr 340
QY 6526 TCTTCTTGTGGCTATCTAGCTTCGGGCCCACTTACTATGAAGGAATGGCTAGAAGA 6585
Db 341 SerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTrpGlyMetAlaArgArg 360
QY 6586 GGGAAATTCATGTGACAAAAGAACATAGAGACCAATGCATGGGATCCCAAAATAAG 6645
Db 361 GlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLys 380
QY 6646 CTTACCTTACTAGGTTCTCGAAAAGGCACCTGCATAGGAAGTTCCCCCATCCAC 6705
Db 381 LeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHis 400
QY 6706 CAACACCTTGTAAACACACTGAAGCTTTAATCAAACTCTCAGAGTCAATATCGTA 6765
Db 401 GlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnThrLeuVal 420
QY 6766 CTTGGTTATGACAGGTGGTGGCATGTAATPACTGGATTAAACCCCTTGTGTTCCACCTTG 6825
Db 421 ProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeu 440
QY 6826 GTTTTAAACCAACTAAAGATTTTGCATTATGTCGAATGTCCTCCGAGTGATTAC 6885
Db 441 ValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrTr 460
QY 6886 TATCCCGAAAAGCAATCCTTGATGAATAGTACAGAAATCATCGACAAAAGAGAGAA 6945
Db 461 TyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLysArgGlu 480
QY 6946 CCATATCTCTGACACTTGTGTGATGCTCGGACTTGGAGTGGCAGAGTGTAGGAACA 7005
Db 481 ProfileSerLeuThrLeuAlaValMetLeuLeuGlyValAlaAlaGlyValGlyThr 500
QY 7006 GGAACAGCTGCCCTGTGTCAGGACACACAGCAGCTAGAAACAGGACTTAGTACCTACAT 7065
Db 501 GlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHis 520
QY 7066 CGAATTGTACAGAAAGTCTCAAGCCCTAGAAAATCTGTCAAGTAACTGGAGGAAATCC 7125
Db 521 ArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluSer 540
QY 7126 CTAACCTCTTATCTGAAGTAGTCTTACAGAATAGAGAGGGTTAGATTATTATTCTA 7185
Db 541 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 560
QY 7186 AAGAAGAGGAGTATTGTGTAGCTTGAAGGAGGAATGCTGTTTTTATGTGATCATTTA 7245
Db 561 LysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAspHisSer 580
QY 7246 GGGGCCATCAGAGCTCCATGAACAGCTTAGAGAAAGTTGGAGAGCGTCCAGGGAA 7305
Db 581 GlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArgGlu 600

QY 7306 AAGAAACTACTCAAGGGTGGTTTGAGGGATGTTCAACAGAGTCTCTTTTGGTGGCTACC 7365
Db 601 LysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeuAlaThr 620
QY 7366 CTACTTCTGCTTTAACAGGACCCCTTAATAGTCTCTCTCTGTTACTCACAGTTGGCCA 7425
Db 621 LeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrValGlyPro 640
QY 7426 TGTATTATTAAACAAGTTAATTGCTTCTATTAGAGAAGTAATAAGTCAGTCCAGATCATG 7485
Db 641 CysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMet 660
QY 7486 GTACTTAGACAAAGTACCAAGCCCTGTAGCAGGAAAGCTGGCCGTAGTCTTAC 7542
Db 661 ValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg---LeuTyr 678
RESULT 9
US-10-441-949-38
; Sequence 38, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-38
Alignment Scores:
Prod. No.: 1.8e-212 Length: 678
Score: 3318.50 Matches: 636
Percent Similarity: 93.8% Conservative: 1
Best Local Similarity: 93.7% Mismatches: 3
Query Match: 22.7% Indels: 39
DB: 4 Gaps: 8
US-10-723-552-3 (1-8132) x US-10-441-949-38 (1-678)
QY 5620 ATGCATCCACAGCTTAAACCGGCGCCACCTCCCGATTCCGGGTGGAAAGCGGAAAGACTG 5679
Db 1 MethisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGGTGCTTCTTACTCTGTCAATACTCTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 5740 ACTAATGGTATCGCATAGGACAGCCTGAACCTCCATAAACCCCTTATCTCTCACCTGG 5799
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 5800 TTAATTACTGACTCCGGGACAGGTATTATATCAACACACTCAAGGGGAGGCTCCTTTA 5859
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 5860 GGAACTGTGTGCTGATCTATACCTTTCCTTCAGATCAGTATTCTTCTAGTCTG----- 5913
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuA-gserValIleProSerLeuAsnAsp 100
QY 5914 ---ACCTACCCCCAGATATCTCCATGCTCAGGATTTTATGTTGTTGCCGACGACCACCA 5970

```
Db 101 GlnThrSerProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProPro 120
QY 5971 AATAATGGAACAACTTCGGGAATCCAGAGATTCTTTTGTAAACAATGGAACTGTGTA 6030
Db 121 AsnAsnGlyIysHisCysGlyAsnProArgAspPhePheCysIysGlnTrpAsnCysVal 140
QY 6031 ACCTCTAATGATGATATTCGAATGCGCAACCTCTCAGCAGATAGGGTAAGTTTCT 6090
Db 141 ThrSerAsnAspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSer 160
QY 6091 TATGTCACACCTATACACAGCTCTGGCAATTTAATATAC-----CTGACC 6135
Db 161 TyrValAsnThrTyrThrSerSerGlyGlnPheAsnTyrGlyHisGlyAArgTrpLeuThr 180
QY 6136 TGG-----ATTAGAACTGGAAGCCCCAAGTGTCTCTCCTTCAGAC 6174
Db 181 TrpGlnGlnArgValGlnIysAspIleArgThrGlySerProLysCysSerProSerAsp 200
QY 6175 CTAGATTACCTAAATAAGTTTCACCTGAGAAAGGAAAAACAAGAAATATCCTAAATGG 6234
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyGlnGluAsnIleLeuLysTrp 220
QY 6235 GTAATGCTATGCTTCGGGAAATGGTATATTATGAGGCTCGGGTAAACAACAGGCTCC 6294
Db 221 ValAsnGlyMetSerTrpGlyMetValTyrTyrGlySerGlyIysGlnProGlySer 240
QY 6295 ATTCTAACTATTCGCTCAAAATAAAC---CAGCTGAGGCTCCAATGGCTATAGGACCA 6351
Db 241 IleLeuThrIleArgLeuLysIleAsnThrGlnLeuGluProProMetAlaIleGlyPro 260
QY 6352 AATACGCTCTGAGGCTCAAGACCCCAACCCCAAGGACCA----- 6393
Db 261 AsnThrValLeuThrGlyGlnArgProProThrGlnGlyProProHisAsnLeuProVal 280
QY 6394 -----GGACCATCTCTCT-----AACATAACTTCTGGA 6420
Db 281 ProGlnGlyProSerProAsnProAspIleThrGlnSerAspTyrAsnIleThrSerGly 300
QY 6421 TCAGACCCCACT-----GAGTCTAACAGCAGCTAAATGGGGGCAAAA 6465
Db 301 SerAspProThrAsnThrProArgAsnGluSerAsnSerThrThrLysMetGlyAlaLys 320
QY 6466 CTTTATTAGCTCATCCAGGAGCTTTTCAAGCTTTTAAGCTTCAAGCTCCAGACTCCAGGCTACC 6525
Db 321 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThr 340
QY 6526 TCTTCTTTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAAGA 6585
Db 341 SerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAlaArgArg 360
QY 6586 GGGAAATTCATGTGACAAAGAACATAGACACCAATGCATGGGGATCCCAAAATAAG 6645
Db 361 GlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLys 380
QY 6646 CTTACCTTACTAGGTTTCTGGAAAAGCCCTGCATAGGAAGGTTCCCCCATCCAC 6705
Db 381 LeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHis 400
QY 6706 CAACACCTTTGTACACCACTGAAGCTTTAATCAAAACCTCTGAGAGTCAATATCTGGTA 6765
Db 401 GlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyrLeuVal 420
QY 6766 CCTGGTTATGACAGGCTGGGCATGTAATACCTGATTAACCCCTGTGTTCACCTTG 6825
Db 421 ProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeu 440
QY 6826 GTTTTAAACCAACTAAGATTTTGCATTATGCTCCAAATGTTCCCGGAGTGATTAC 6885
Db 441 ValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrTrp 460
QY 6886 TATCCCGAAAAGCAATCCTTTGATGAATATGACTTACAGAAATCATCGACAAAAGAGAA 6945
```

```
Db 461 TyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLysArgGlu 480
QY 6946 CCATATCTCTGCACACTGCTGTGATGCTCGGACTTTGGAGTGCAGCAGGTGTAGGAACA 7005
Db 481 ProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThr 500
QY 7006 GGAACAGCTGCGCTGTCTACCGGACCACAGCAGCTAGAAACAGGACTTAGTAACCTACAT 7065
Db 501 GlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHis 520
QY 7066 CGAATTGTAAACAAGATCTCCAAGCCCTAGAAAAATCTGTCAAGTAACCTGGAGGAATCC 7125
Db 521 ArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluSer 540
QY 7126 CTAACTCCTTATCTGAAGTAGTCTTACAGAATAGAGAGGGTTAGATTTATTATTCTA 7185
Db 541 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 560
QY 7186 AAGAAGGAGGATTATGTGTAGCTTTGAAGGAGGAATGCTGTTTTTATGTGGATCATTC 7245
Db 561 LysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAspHisSer 580
QY 7246 GGGGCCATCAGAGACTCCATGAACAAGCTTTAGAAAGGTTGGAGAAGCGTCCGAAGGAA 7305
Db 581 GlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArgGlu 600
QY 7306 AAGAAACTACTCAAGGGTGGTTGAGGATGCTTCAACAGGCTCTTTGGTTGGCTACC 7365
Db 601 LysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeuAlaThr 620
QY 7366 CTACTTTCTGCTTTAAACAGGACCTTTAAVAGTCTCTCTCTGTACTCACAGTTGGGCCA 7425
Db 621 LeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrValGlyPro 640
QY 7426 TGTATTATTAAACAAGTTAATTGCCTTCATTAGAAACGAATAAGTGCAGTCCAGATCATG 7485
Db 641 CysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMet 660
QY 7486 GTACTTAGACAACAGTACCAAGCCGCTTAGCAGGAGGAGCTGGCCGTAGCTCTAC 7542
Db 661 ValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg---LeuTyr 678
```

RESULT 10
US-10-029-656-2
; Sequence 2, Application US/10029656
; Publication No. US20030175904A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; APPLICANT: Oldmixon, Beth
; APPLICANT: Ericsson, Thomas
; TITLE OF INVENTION: Molecular Sequence of Pig Endogenous Retrovirus Receptor and Meth
; FILE REFERENCE: 329579-3
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/10/029,656
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Viral
; US-10-029-656-2

Alignment Scores:
Pred. No.: 4.64e-201 Length: 653
Score: 3147.50 Matches: 593
Percent Similarity: 93.6% Conservative: 18
Best Local Similarity: 90.8% Mismatches: 27
Query Match: 21.5% Indels: 15
DB: 4 Gaps: 3

US-10-723-552-3 (1-8132) x US-10-029-656-2 (1-653)

QY 5620 ATGCATCCCAAGTAAACCGCGCCACCTCCCGATTCGGGTGGAAGCGAAGAAAGACTG 5679
Db : : : : :
1 MethisProThrLeuSerArgArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTCATAAACCCTCTCAG 5739
Db : : : : :
21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 5740 ACTAATGGTATCGCATAGGACAGCAGCTGAACCTCCCATAAACCTTATCTCCACCTGG 5799
Db : : : : :
41 ValAsnGlyLysArgLeuValAsnSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 5800 TTAATTAATCACTCCGCGCACAGGTATTAATATCAACAACAACACTCAAGGGGAGGCTCTCTTA 5859
Db : : : : :
61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 5860 GGAACCTGGTGGCTGATCTATAGTTGGCTCAGATCAGTATTCCTAGTCTG----- 5913
Db : : : : :
81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 5914 ---ACCTCACCCACAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCAACA 5970
Db : : : : :
101 GlnAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
QY 5971 AATAATGGAAACAATTCGGAAATCCAGAGATTTCTTTTGTAAACAATGGAACCTGTGTA 6030
Db : : : : :
121 AsnAsnGluTyrCysGlyAsnProGlnAspPheCysLysGlnTrpSerCysVal 140
QY 6031 ACCTTAATGATGATATGGAAATGGCCAACTCTCAGCAGGATAGGGTAAGTTTCT 6090
Db : : : : :
141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY 6091 TATGTCAACACCTATACCACTCGGACATTAATTAAC-----CTGACC 6135
Db : : : : :
161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY 6136 TGG-----ATTAGAACTGGAAAGCCCAAGTCTCTCTCCAGAC 6174
Db : : : : :
181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 6175 CTAGATTACTTAAATAAGTTTCACTGGAAGAGGAAACAAGAAATATCTTAAATGG 6234
Db : : : : :
201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlnAsnIleLeuLysTrp 220
QY 6235 GTAATAGGTATGCTTGGGAATGGTATATTATGAGGCTCGGTTAAACAACAGGCTCC 6294
Db : : : : :
221 ValHisGlyMetSerTrpGlyMetValTyrTyrGlyLysGlyLysGlnProGlySer 240
QY 6295 ATTCTAACTATTCCCTCAAAATAAACCACTCGAGCTCCCAATGGCTATAGGACCAAAAT 6354
Db : : : : :
241 IleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMetAlaIleGlyProAsn 260
QY 6355 ACGGCTTGACGGGTCAAGACCCCAACCCAAAGGACCGAGGACCATCTCTAATAACT 6414
Db : : : : :
261 ThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyThrSerSerAsnIleThr 280
QY 6415 TCTGGATCAGACCCCACTAGTCTAACACGACCACTAAATGGGGCAAACTTTTACG 6474
Db : : : : :
281 SerGlySerAspProThrGluSerAsnSerThrThrLysMetGlyAlaLysLeuPheSer 300
QY 6475 CTATCCAGGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAGGCTACCTCTTCTGT 6534
Db : : : : :
301 LeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThrSerSerCys 320
QY 6535 TGCTATGCTTACGCTTGGGCCCACTTATATGAAGGAATGGCTAGAGAGGAAATTC 6594
Db : : : : :
321 TrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAlaArgGlyLysPhe 340
QY 6595 AATGTGACAAAGAACATAGACCAATGACATGGGGATCCCAAAATAGCTTACCTT 6654
Db : : : : :
341 AsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLysLeuThrLeu 360

QY 6655 ACTGAGGTTTCTGAAAAAGGCACCTGCATAGGAAAGGTTCCCCATCCCAACCAACACTT 6714
Db : : : : :
361 ThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu 380
QY 6715 TGTAAACCACTGAAGCTTTAATCAAACTCTGAGAGTCAATATCTGGTACCTGGTTAT 6774
Db : : : : :
381 CysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyrLeuValProGlyTyr 400
QY 6775 GACAGGTGGTGGCATGTAATACCTGGAATTAACCCCTTGTCTTCCACTTGGTTTAAAC 6834
Db : : : : :
401 AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn 420
QY 6835 CAAACTAAAGATTTTGCATTTATGTCCTCAAAATTTGTTCCCGAGTGTATTAATCTCCGAA 6894
Db : : : : :
421 GlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrTyrProGlu 440
QY 6895 AAGCAATCTTCATCAATATGACTACAGAAATCATCGAAGAGAGAGAACCCATATCT 6954
Db : : : : :
441 LysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLysArgGluProIleSer 460
QY 6955 CTGACACTTGTCTGATGCTCGGACTTGGAGTGGCAGCAGGTGTAGGAAACAGAAACAGCT 7014
Db : : : : :
461 LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGlyValGlyThrGlyThrAla 480
QY 7015 GCCTCTGTCACGGGACACAGCAGCTAGAAAACAGGACTTAGTAACTCATCTCGAATTGTA 7074
Db : : : : :
481 AlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHisArgIleVal 500
QY 7075 ACGAAGATCTCAAGCCCTAGAAAAATCTGTCACTAACTCGAGGAATCCCTAACTCC 7134
Db : : : : :
501 ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSerLeuThrSer 520
QY 7135 TTATCTGAAGTAGTCTCAGAAATAGAGGGTGTAGATTATTATTCTTAAAGAGGA 7194
Db : : : : :
521 LeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeuPheLeuLysGly 540
QY 7195 GGATATTGTAGTACCTTGAAGGAGGAATGCTGTTTATTGTGATCATTCAGGGGCATC 7254
Db : : : : :
541 GlyLeuCysValAlaLeuLysGluGluCysPheTyrValAspHisSerGlyAlaIle 560
QY 7255 AGAGACTCATGAACAGCTTAGAGAAAGGTTGGAGAACGCTGGAAGGGAAGAAAGAACT 7314
Db : : : : :
561 ArgAspSerMetSerLysLeuArgGluArgLeuGluLysArgArgGluLysGluThr 580
QY 7315 ACTCAAGGCTGTTGAGGGATGGTTCAACAGTCTCTTTGTTGGCTACCTACTTCT 7374
Db : : : : :
581 ThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeuAlaThrLeuLeuSer 600
QY 7375 GCTTTAAACAGGACCCCTTAATAGTCTCTCTCTTACTCACAGTTGGCCCATCTATTATT 7434
Db : : : : :
601 AlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrValGlyProCysIleIle 620
QY 7435 AACAGTTAATTCCTTCATTATAGAGAACGAATAAGTGCAGTCCAGATCATGTGTTAGTA 7494
Db : : : : :
621 AsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMetValLeuArg 640
QY 7495 CAACAGTACCAAGCCGCTAGCAGGGAAGCTGGCCGC 7533
Db : : : : :
641 GlnGlnTyrGlnSerProSerArgGluAlaGlyArg 653

RESULT 11

US-10-029-656-4
; Sequence 4, Application US/10029656
; Publication No. US20030175904A1
; GENERAL INFORMATION:
; APPLICANT: Clive
; APPLICANT: Oldmixon, Beth
; APPLICANT: Erickson, Thomas
; TITLE OF INVENTION: Molecular Sequence of Pig Endogenous Retrovirus Receptor and Met
; TITLE OF INVENTION: Use
; FILE REFERENCE: 329579-3
; CURRENT APPLICATION NUMBER: US/10/029,656

; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/60/285,103
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Viral
US-10-029-656-4

Alignment Scores:

Pred. No.: 2,94e-200 Length: 653
Score: 3135.50 Matches: 591
Percent Similarity: 93.1% Conservatives: 17
Best Local Similarity: 90.5% Mismatches: 30
Query Match: 21.4% Indels: 15
DB: 4 Gaps: 3

US-10-723-552-3 (1-8132) x US-10-029-656-4 (1-653)

Qy	5620	ATGCATCCCAAGTAAACCGCGCCACCTCCGATTCGGGTGGAAACCGAAAGACTG	5679
Db	1	MetHisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu	20
Qy	5680	AAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTCATAAATCTCAG	5739
Db	21	LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln	40
Qy	5740	ACTAATGGTATGCGATAGAGACAGCTCAACTCCCATAAACCTTATCTCTCACCTGG	5799
Db	41	ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp	60
Qy	5800	TTAATTAAGTACCTCCGACAGAGTATTAATATCAACAACACTCAAGGGAGGCTCTTTA	5859
Db	61	LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu	80
Qy	5860	GGAACTGGTGGCTGATCTATAGCTTTGGCTCAGATCAGTTATCTCTAGTCTG	5913
Db	81	GlyThrTrpProGluLeuTyValCysLeuArgSerValIleProGlyLeuAsnAsp	100
Qy	5914	---ACCTCACCCAGATATCTCTCATCTCAGCTCAGGATTTATGTTCCCGACGACCA	5970
Db	101	GlnAlaThrProProAspValLeuArgAlaTyGlyPheTyValCysProGlyProPro	120
Qy	5971	ANTAATGGAAACATCCGAAATCCAGAGATTTCTTTGTAAACAATGGAATGTGA	6030
Db	121	AsnAsnGluGluTyCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal	140
Qy	6031	ACCTTAATGATGGATATTGGAATGGCCACCTCTCAGCAGGATAGGTAAGTTTCT	6090
Db	141	ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTySer	160
Qy	6091	TATGTCAACCTATACAGCTCTGGACAATTTAATTAC-----CTGACC	6135
Db	161	PheValAsnAsnProThrSerTyAsnGlnPheAsnTyGlyHisGlyArgTrpLysAsp	180
Qy	6136	TGG-----ATTAGAACTGGAAAGCCCAAGCTCTCTCTTCAGAC	6174
Db	181	TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp	200
Qy	6175	CTAGATTACCTAAATAAGTTTCACTGAGAAAGGAAACAAAGAAATATCTTAAATGG	6234
Db	201	LeuAspTyLeuLysIleSerPheThrGluLysGlyLysGlnGlnAsnIleLeuLysTrp	220
Qy	6235	GTAATGGTATGCTTCGGGAAATGGTATATTATCGAGCTCGGTAAACAACAGGCTCC	6294
Db	221	ValAsnGlyMetSerTrpGlyMetValTyTyGlySerGlyLysGlnProGlySer	240
Qy	6295	ATTCTAACTATTCGCTCAAAATAAACAGCTGGAGCTCCAATGGCTATAGGACCAAT	6354
Db	241	IleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMetAlaIleGlyProAsn	260

Qy	6355	ACGGTCTTGACGGTCAAGACCCCAACCCAGGACCCAGGACCATCTCTTAACATAACT	6414
Db	261	ThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyProSerSerAsnIleThr	280
Qy	6415	TCTGGATCAGACCCCACTGAGTCTAAACAGCACCACTAAAAATGGGGGCAAACTTTT	6474
Db	281	SerGlySerAspProThrGluSerAsnSerThrThrLysMetGlyAlaLysLeuPheSer	300
Qy	6475	CTCATCCAGGAGCTTTTCAAGCTCTTAACCTCACGACTCCAGAGCTACTCTCTTTGT	6534
Db	301	LeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlnAlaThrSerSerCys	320
Qy	6535	TGGCTATGCTTAGCTTCGGGCCCCACCTTACTATGAAGGAATGGCTAGAGAGGAAATC	6594
Db	321	TrpLeuCysLeuAlaLeuGlyProProTyTyGlyMetAlaArgArgGlyLysPhe	340
Qy	6595	AATGTGACAAAAGAACATAGAGACCAATGCACATGGGGATCCCAAAATAAGCTTTAC	6654
Db	341	AsnValThrLysGlyHisArgAspProCysThrTrpGlySerGlnAsnLysLeuThrLeu	360
Qy	6655	ACTGAGTTTCTGAAAGGACCTGCATAGAGAAAGTTTCCCATCCCAACCAACCTT	6714
Db	361	ThrGluValPheGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu	380
Qy	6715	TGTAACCAACACTGAAGCTTTAATCAAACTCTGAGAGTCAATATCTGGTACCTGGTAT	6774
Db	381	CysAsnHisThrGluAlaPheAsnArgThrSerGluSerGlnTyLeuValProGlyTy	400
Qy	6775	GACAGTGGTGGGCATGTAATCTGGAATTAACCCCTTGTGTTCCTCCACTTGGTTTTAAC	6834
Db	401	AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn	420
Qy	6835	CAAACTAAAGATTTTTCATATTATGTCCTCAAAATTTGTCCTCCGAGTGATTTACTAT	6894
Db	421	GlnThrLysAspPheCysIleMetValGlnIleValProArgValTyTyTyTyProGlu	440
Qy	6895	AAAGCAATCTTCATCAATATGACTACAGAAATCATCGAACAAAGAGAGAACCATATCT	6954
Db	441	LysAlaIleLeuAspGluTyAspTyArgAsnHisArgGlnLysArgGluProIleSer	460
Qy	6955	CTCACACTTGTCTGTGCTCGGACTTGGAGTGGCAGCAGTGCTAGGAAACAGGAACAGCT	7014
Db	461	LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThrGlyAla	480
Qy	7015	GCCTGTGTCGGGACACAGCAGCTAGAAACAGGACTTAGTAACCTACATCGAATTGTA	7074
Db	481	AlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHisArgIleVal	500
Qy	7075	ACAGAAAGATCTCCAAGCCCTAGAAAAATCTGCAGTAACCTGGAGGAATCCCTAACCTCC	7134
Db	501	ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSerLeuThrSer	520
Qy	7135	TTATCTGAAGTAGTCTTACAGAAATAGAGAGGTTAGATTATTATTCTTAAAGAAAGGA	7194
Db	521	LeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuLysGluGly	540
Qy	7195	GGATTATGTAGCTTGAAGGAGGAATGCTGTTTTTATGTGATCATTCAGGGGCCATC	7254
Db	541	GlyLeuCysValAlaLeuLysGluGlyCysPheTyValAspHisSerGlyAlaIle	560
Qy	7255	AGAGACTCCATGAACCAAGCTTAGAGAAAGGTTGAGAGAGCGTCGAAGGGAAGGAACCT	7314
Db	561	ArgAspSerMetSerLysLeuArgGluArgLysArgArgGluLysGluThr	580
Qy	7315	ACTCAAGGGTGGTTTGAAGGATGGTTCAACAGGTCTCTTTTGGTTGGCTACCTACTTCT	7374
Db	581	ThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeuAlaThrLeuLeuSer	600
Qy	7375	GCTTTAACAGGACCTTAAATAGTCTCTCTCTGTTACTCAGCTTGGGCCATGTATTATT	7434
Db	601	AlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrValGlyProCysIleIle	620
Qy	7435	AACAAGTTAATTCCTTATTAGAGAACGAATAAGTGCACTCCAGATCATGGTACTTAGA	7494


```
Db 621 AsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMetValLeuArg 640
QY 7495 CAACAGTACCACAGCCCGTCTAGCAGGGAAGCTGGCCGC 7533
Db 641 GlnGlnTyroGlnSerProSerSerArgGluAlaGlyArg 653

RESULT 12
US-10-441-949-43
; Sequence 43, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; PRIOR FILING DATE: 2003-05-19
; PRIOR FILING DATE: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-43

Alignment Scores:
Pred. No.: 3 65e-183 Length: 661
Score: 2879.50 Matches: 541
Percent Similarity: 88.3% Conservative: 40
Best Local Similarity: 82.2% Mismatches: 56
Query Match: 19.7% Indels: 21
DB: 4 Gaps: 6

US-10-723-552-3 (1-8132) x US-10-441-949-43 (1-661)
QY 5620 ATGCATCCACGTTAAACCGGCCACCTCCCGATTCCGGGTGGAAGCGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCGCTCCATCGGTGTTCTTACTCTGTCAATACCTCTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 5740 ACTAATGGTATGGCATAGGACACCTGAACCTCCATAAACCTTATCTCTCACCTGS 5799
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 5800 TTAATTAAGTACCTCCGGCACAGGTATTAATATCAACACACTCAAGGGGAGGCTCTTTA 5859
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 5860 GGAACTGGTGGCTGATCTATACGTTTGCTCAGATCAGTATTATTCCTAGTCTG----- 5913
Db 81 GlyThrTrpProGlnLeuTyroValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 5914 ---ACCTCACCCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGCCACCA 5970
Db 101 GlnAlaThrProProAspValLeuArgAlaTyroGlyPheTyroValCysProGlyProPro 120
QY 5971 AATAATGGAACCAATTCGGAATCCAGAGATTTCTTTTGTAAACAATGGAAGTCTGTA 6030
Db 121 AsnAsnGluTyroCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 6031 ACCTCTAATGATGATATTGGAAATGGCCAACTCTCAGCAGGATAGGTTATTTTCT 6090
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTrpSer 160
QY 6091 TATGTCAACACCTATACAGCTCTCGACATTTAAATTAC-----CTGACC 6135
Db 161 PheValAsnAsnProThrSerTyroAsnGlnPheAsnTyroGlyHisGlyArgTrpLysAsp 180
QY 6136 TGG-----ATTAGAACCTGGGAAGCCCCAGTGTCTCTCTCTTCAGAC 6174
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysAsnSerLeuAsp 200
QY 6175 CTAGATTACTTAAATAAGTTTCACTGAGAAAGGAAGAAACAAGAAATATCTTAAATGG 6234
Db 201 LeuAspTyroLysIleSerPheThrGluLysGlyGlnGlnAsnIleGlnLysTrp 220
QY 6235 GTAAATGGTATGTTGGGAATGGTATATTATGGAGGCTCGGTAAACAACACAGGCTCC 6294
Db 221 ValAsnGlyMetSerTrpGlyIleValTyroGlyLysSerGlyArgLysLysGlySer 240
QY 6295 ATTCTAACTATTGCGCTCAAAATA---AACCAGCTGGAGCTCCAAATGGCTATAGACCA 6351
Db 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProValAlaIleGlyPro 260
QY 6352 AATACGGTCTTGACGGGTCAAGACCCCAACCCAA-----GGACCAGGACCA 6399
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGlnArgProSerProAsnPro 280
QY 6400 TCCTCT---AACATAACTTCTGGATCAGACCCCACTGAGTCTAACACGACGACTAAATG 6456
Db 281 SerAspTyroAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY 6457 GGGGCAAACTTTTACGCTCATCCAGGGAGCTTTTCAAGCTCTTAACTCCAGACTCCA 6516
Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY 6517 GAGGCTACCTCTTCTTGGTGTAGCTTACGTTCCGGGCCACCTTACTATGAGGAATG 6576
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyroTyroGlyMet 340
QY 6577 GCTAGAAGAGGGAATTCATATGCAAAAGAACATAGAGACCAATGCATCGCATGGGATCC 6636
Db 341 AlaArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 6637 CAAATAACCTTACCTTACGAGGTTCTGGAAGAGGACCTGCATAGAGAAAGTTTCCC 6696
Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyArgValPro 380
QY 6697 CATCCCAACACACCTTTGTAACACACTGAAGCCCTTAAATCAAACTCTGAGAGTCAA 6756
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 6757 TATCTGGTACCTGGTTTATGACAGGTGGTGGCATGTAATCTGGATTAAACCTCTGTGT 6816
Db 401 TyrLeuValProGlyTyroArgTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 6817 TCCACCTGGTTTAAACCAAACTAAAGATTTTTCATTATGTCCTCAAAATGTTCTCCCGA 6876
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 6877 GTGTATTACTATCCGGAAGAAAGCAATCCTTGATGATATGACTACAGAAATCATCCACA 6936
Db 441 ValTyroTyroProGluLysAlaValLeuAspGluTyroAspTyroArgTrpAsnArgPro 460
QY 6937 AAGAGAACCCATATCTCTGACACTTGTGTGATGCTCGGACTTCGAGTGGCAGCAGGT 6996
Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGly 480
QY 6997 GTAGGAACAGGAACAGCTGCCCTGGTACGGGACCAACAGCAGCTAGAAACAGGACTTAGT 7056
Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
QY 7057 AACCTACATCGAATTTGTAACAGAAATCTCCAAAGCCCTAGAAAAATCTCTCAGTACCTG 7116
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
```

```

QY 7117 GAGGAATCCCTAACCTCTTATCTGAAGTAGTCTTACAGAAATAGAGAGGTTAGATTTA 7176
Db 521 GUGIUSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 7177 TTATTTCTAAAGAGAGGAGATTATGTGTAGCTTTGAAGGAGGAATCTGTTTTTATGTG 7236
Db 541 LeuPheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal 560
QY 7237 GATCATTGAGGGCCATCAGAGACTCCATGAACAAAGCTTGAAGAAAGGTTGGAGACGCT 7296
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArg 580
QY 7297 CGAAGGGAAGGAACTACTCAAGGCTGTTTTCAGGAGTGGTTCAACAGGCTCTTTGG 7356
Db 581 ArgArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyr 600
QY 7357 TTGGCTACCTTCTGCTTTTAAACAGGACCTTAAATAGTCTCTCTCTTACTCACA 7416
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 620
QY 7417 GTTGGCCATGTATTATTAACAAGTTAATTGCGCTTCAATTAAGAGAACAAATAGTGCAGTC 7476
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
QY 7477 CAGATCATGTTACTAGACACACTACCAAGCCGCTCTAGCAGGGAAGCTGGC 7530
Db 641 GlnIleMetValLeuArgGlnGlnTyrGlnGlyLeuLeuSerGlnGlyGluGly 658

```

RESULT 13

US-10-441-949-16

; Sequence 16, Application US/10441949

; Publication No. US20040116684A1

; GENERAL INFORMATION:

; APPLICANT: Rodzigo, Allen

; APPLICANT: Ross, Howard A.

; APPLICANT: Mullins, James I.

; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES

; FILE REFERENCE: 08987-012001

; CURRENT APPLICATION NUMBER: US/10/441,949

; CURRENT FILING DATE: 2003-05-19

; PRIOR APPLICATION NUMBER: PCT/US01/05288

; PRIOR FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: US 60/183,659

; PRIOR FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 660

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Artificially generated peptide

US-10-441-949-16

```

Alignment Scores:
Pred. No.: 7,86e-183 Length: 660
Score: 2874.50 Matches: 539
Percent Similarity: 88.8% Conservative: 38
Best Local Similarity: 82.9% Mismatches: 52
Query Match: 19.6% Indels: 21
DB: 4 Gaps: 6

```

US-10-723-552-3 (1-8132) x US-10-441-949-16 (1-660)

```

QY 5620 ATGCATCCCAAGTTAAACCGCGCACCTCCGATTCGGGTGGAAAGCGGAAAGACTG 5679
Db 1 MethHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 5680 AAAATCCCTTAAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40

```

```

QY 5740 ACTAATGGTATGCGCATAGGAGACAGCCTGAACTCCATCAAAACCCCTTATCTCACCTGG 5799
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTyr 60
QY 5800 TTAATTAAGTACTCCGCGCAGAGTATTAAATATCAACAACACTCAAGGGGAGGCTCCTTTA 5859
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 5860 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTATTCTAGTCTG----- 5913
Db 81 GlyThrTyrTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 5914 ---ACCTCACCCCCAGATATCTCCATCTCAGGATTTTATGTTGCCAGAGACCAACCA 5970
Db 101 GlnAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
QY 5971 AATAATGGAACCAATTCGCGAAATCCAGAGATTTCTTTTGTAAACAATGGAATCTGTGTA 6030
Db 121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 6031 ACCTCTAATGATGATATTGGAATTCGCAACCTCTCAGCAGATAGGTAAGTTTCT 6090
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY 6091 TATGTCAACACCTATACAGCTCTGGACAAATTTAATTAC-----CTGACC 6135
Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY 6136 TGG-----ATTAGAACTGGAAGCCCCCAAGTCTCTCTCTTCACAG 6174
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 6175 CTAGATTACCTAAATAAGTTTCACCTAGAGAAAGGAAACAAAGAAATATCTTAAGTGG 6234
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTyr 220
QY 6235 GTAAATGGTATGTCTTGGGAATGCTATATTATGGAGGCTCGGTAACAAACACAGGCTCC 6294
Db 221 ValAsnGlyMetSerTrpGlyIleValTyrGlySerGlyArgLysLysGlySer 240
QY 6295 ATTCTAATTAATCCCTCAAAATA---AACCAGCTGGAGCTCCAATGGCTATAGGACCA 6351
Db 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProValAlaIleGlyPro 260
QY 6352 AATACGGTCTTGGAGGTCGCAACAGCCCCCAACCCAA-----GGACCCAGGACCA 6399
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGlnArgProSerProAsnPro 280
QY 6400 TCCTCT---AACATAACTCTGGATCAGACCCCACTGAGTCTAACAGCAGCAGCTAAATG 6456
Db 281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY 6457 GGGGCAAAACCTTTTACGCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCAGACTCCA 6516
Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY 6517 GAGCTACCTCTCTTGTGGCTATGCTTACGTTCCGGCCCACTTACTTATGAGGAATG 6576
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMet 340
QY 6577 GCTAGAGAGGGAAATTTCAATGTGACAAAGAACATAGACCAATGACATGGGATCC 6636
Db 341 AlaArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 6637 CAAAATAAGCTTACCTTACTGAGGTTTCTGGAAGAGCCCTGTCATPAGAAAGGTTCCC 6696
Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyArgValPro 380
QY 6697 CATCCCAACCAACCTTTGTAAACACACTGAAGCCTTTAATCAACCTCTGAGAGTCAA 6756
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 6757 TATCTGTGCTACCTGTTATGACAGTGTGTGGCATGTAATACTGGATTAAACCCCTTGTGTT 6816

```

Db 401 TyLeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 6817 TCCACCTTGGTTTTAAACCAAACTAAAGATTTTGGCAATATGTCCTCAAAATGTTCCCGCA 6876
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 6877 GTGATTACTATCCCGAAAAAGCAATCTTGTGATGAATATGATACACAGAAATATCATCGACAA 6936
Db 441 ValTyTyrProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460
QY 6937 AAGAGAGAACCCATATCTGACACTGCTGTGATGCTCGGACTTGGAGTGGCAGCAGGT 6996
Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
QY 6997 GTAGGAACAGAACAGCTGCGCTTGGTCACGGGACACAGCAGCTAGAAAACAGACTTAGT 7056
Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGlyGlyLeuSer 500
QY 7057 AACCTACATCGAATGTAAACAGAGATCTCCAAAGCCCTAGAAAAATCTGTGAGTAACCTG 7116
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 7117 GAGGAATCCCTAACCTCTTATCTGAGTAGTCTCTACAGATACAGAGAGGTTAGATTTA 7176
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 7177 TTATTTCTAAAGAGAGGAGGATTATGTAGCTTGAAGGAGGAATGCTGTTTTATG 7236
Db 541 LeuPheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal 560
QY 7237 GATCATTCAGGGCCCATCAGAGACTCCATGAACAGCTTAGAGAAAGTTGGAGAGCGT 7296
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 580
QY 7297 CGNAGGAAAGAAACTACTCAAGGTGTTTGGGGATGTTCAACAGGCTCTTTGG 7356
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 7357 TTGGCTACCTACTTCTGCTTTAAGCAGGACCTTAATAGTCTCTCTCTGTTACTCACA 7416
Db 601 MetThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThr 620
QY 7417 GTGGCCCATGTATTATTAACAGTTAATGCTTCAATAGAGAACGAATAGTCAGTC 7476
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
QY 7477 CAGATCATGCTACTTAGACACAGTACCAA 7506
Db 641 GlnIleMetValLeuArgGlnGlnTyrGln 650

RESULT 14

US-10-441-949-8
; Sequence 8, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-8

Alignment Scores:
Pred. No.: 1,25e-182 Length: 660
Score: 2871.50 Matches: 539
Percent Similarity: 88.6% Conservative: 37
Best Local Similarity: 82.9% Mismatches: 53
Query Match: 19.6% Indels: 21
DB: 4 Gaps: 6

US-10-723-552-3 (1-8132) x US-10-441-949-8 (1-660)

QY 5620 ATGCATCCCAACGTTAAACCGGCGCCACCTCCCGATTCCGGGTGAAAGCGAAAGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 5680 AAAATCCCTTAAGCTTCCCTCCATCGCGTGGTCTTACTCTGTCAATAAACCCTTCTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 5740 ACTAATGGTATGGCATAGAGACAGCCTGAACCTCCATAAACCCCTTATCTCTCAGCTGG 5799
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 5800 TTAATTAAGTCTCCGGCAGCAGGTATTATATCAACAACACTCAAGGGAGGCTCCTTTA 5859
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyAlaProLeu 80
QY 5860 GGAACCTGGTGGCTGATCTATACGTTTCCTCAGATCAGTTATTTCTAGTCTG----- 5913
Db 81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 5914 ----ACCTCACCCCGACAGATATCTCCATGCTCAGGATTTATGTTGGCCAGACACCA 5970
Db 101 GlnAlaThrProProAspValLeuArgAlaTrpGlyPheTyrValCysProGlyProPro 120
QY 5971 AATAATGGAACATTCGCGAAATCCAGAGATTTCTTTTGTAAACAATGGAAGTGTGA 6030
Db 121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 6031 ACCTCTAATGATGATATTGGAATGGCCAACTCTCAGCAGGATAGGGTAAGTTTCT 6090
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY 6091 TATGTCACACCTATACAGCTCTGCAATTAATTAC-----CTGACC 6135
Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY 6136 TGG-----ATTAGAACTGGAAGCCCAAGTCTCTCTTCAGAC 6174
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 6175 CTAGATTACCTAAAAATAAGTTTCACTGAGAAAGGAAAAACAAGAAATATCTCTAAATGG 6234
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrp 220
QY 6235 GTAATGATGATGCTTGGGAATGATATTATGAGGCTCGGTAAACAACAGAGCTCC 6294
Db 221 ValAsnGlyMetSerTrpGlyIleValTyrTyrGlyGlySerGlyArgLysGlySer 240
QY 6295 ATTCTAACTATTTCGCTCAAAATA---AACAGCTGGAGCCTCCCAATGCTATAGACCA 6351
Db 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProValAlaIleGlyPro 260
QY 6352 AATACGCTTCTGACGGGTCAAAAGACCCCAACCCAA-----GGACGAGACCA 6399
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGlnArgProSerProAsnPro 280
QY 6400 TCCTCT---AACATACTTCTGGATCAGACCCCACTGAGTCTAACAGCAGCACTAAATG 6456
Db 281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300

```
QY 6457 GGGGCAAACTTTTTCAGCTCATCCAGGAGCTTTTCAAGCTCTTAAGTCCAGACTCCA 6516
Db 301 GLYAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY 6517 GAGGTACCTCTTCTTGTGGCTATGCTTACGTTCCGGCCCACTTACTATGAAGAAATG 6576
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMet 340
QY 6577 GCTAGAGAGGGAAATTCATGTGACAAAGACATAGACCAATGCAATGGGATCC 6636
Db 341 AlaArgGlyGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 6637 CAAAATAGCTTACCTTACGTAGTTCTGGAAAGCGCACCTGCATAGAAAGTTCCC 6696
Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380
QY 6697 CCATCCCAACCAACCTTTGTAAACCACTGAAGCCCTTTAATCAAACTCTGAGAGTCAA 6756
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 6757 TATCTGTACTCTGTTATGACAGGTGGTGGCATGTAAATCTGGATTAACCCCTTGTGTT 6816
Db 401 TyrLeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 6817 TCCACCTTGGTTTAAACCAAACTAAAGATTTTGGCATTATGTCCAAATTTGTTCCCGA 6876
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 6877 GTGTATTACTATCCGAAAAAGCAATCTTGTGATGAATATGACTACAGAAATCATCGACAA 6936
Db 441 ValTyrTyrTrpProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460
QY 6937 AGAGAGAACCATACTCTGACACTGCTGTGATGCTCGGACTTGGAGTGGCAGCAGGT 6996
Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
QY 6997 GTAGGACAGGACAGCTGCTGCTCAGCGGACCCAGCAGCTAGAAACAGGACTAGT 7056
Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
QY 7057 AACCTACATCGAATTGTAACAGAGATCTCAAGCCCTAGAAAAATCTGTCACTAACCTG 7116
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 7117 GAGAAATCCCTAATCTCTTATCTGAAGTAGTCTTACAGAAATGAAAGAGGTAGATTTA 7176
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 7177 TTATTTCTAAAGAGGAGGATATGTGTAGCCCTTGAGGAGGATGCTGCTTTTATGTG 7236
Db 541 LeuPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrVal 560
QY 7237 GATCATTTCAGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAAGGTTGGAGAGCGT 7296
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 580
QY 7297 CGAAGGGAAAAGGAAACTACTCAAGGGTGGTTTGGAGGATGGTTCAACAGGTCTCTTGG 7356
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 7357 TTGGTACCTTCTTCTGCTTTAACAGGACCCCTTAATAGTCTCTCTCTCTTACTCACA 7416
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThr 620
QY 7417 GTTGGGCCATGTATTATTAACAAGATTAAATGCGCTTCAATTAGAGAACCAATAAGTGCAGTC 7476
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
QY 7477 CAGATCATGTGTACTTAGACAACAGTACCAA 7506
Db 641 GlnIleMetValLeuArgGlnGlnTyrGln 650
```

```
RESULT 15
US-09-851-859A-5
; Sequence 5, Application US/09851859A
; Patent No. US20020065407A1
; GENERAL INFORMATION:
; APPLICANT: Banerjee, Papia T.
; APPLICANT: Patience, Clive
; APPLICANT: Andersson, Goran K.
; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 61750-321
; CURRENT APPLICATION NUMBER: US/09/851,859A
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/097,015
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: US 09/376781
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PERV-A
; OTHER INFORMATION: polypeptide sequence taken from GenBank Accession
; OTHER INFORMATION: No. US20020065407A1 Y12238 for comparison.
US-09-851-859A-5

Alignment Scores:
Pred. No.: 2,69e-182 Length: 660
Score: 2866.50 Matches: 537
Percent Similarity: 88.6% Conservative: 39
Best Local Similarity: 82.6% Mismatches: 53
Query Match: 19.6% Indels: 21
DB: 3 Gaps: 6

US-10-723-552-3 (1-8132) x US-09-851-859A-5 (1-660)
QY 5620 ATGCATCCCAAGTAAACCGCGCCACCTCCCGATTCGGGTGGAAAGCCGAAAGACTG 5679
Db 1 MetHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 5680 ARAATCCCTTAAGCTTCGCTCCATCGCTGCTTCTTACTCTGTCAATAAAGCTCTCAG 5739
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 5740 ACTAATGGTATGCGCATAGGAGACAGCTGAACCTCCCATAAACCTTATCTCACCTGG 5799
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 5800 TTAATTAAGTACCTCCGCGACAGAGTATTAATATCAACAACACTCAAGGGAGGCTCCTTTA 5859
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 5860 GGAACCTGGTGGCTGATCTATACGTTGCTCAGATCAGTATTCCTAGCTG----- 5913
Db 81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 5914 ---ACCTCACCCACAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCAACA 5970
Db 101 GluAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
QY 5971 AATAATGGAACAATTCGCGAAATCCAGAGATTTCTTTTGTAAACAAATGGAAGTGTGTA 6030
Db 121 AsnAsnGluGlyTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysIle 140
QY 6031 ACCTCTAATGATGATATTCGAAATGGCAACCTCTCAGCAGGATAGGTAAGTTTCT 6090
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY 6091 TATGTCAACACCTTATACCGACTCTGGACAAATTTAATTAC-----CTGACC 6135
:::|||||
```

```
Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
Qy 6136 TGG-----ATTAGAACTGGAAGCCCAAGTGCCTCTCCTTCAGAC 6174
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
Qy 6175 CTAGATTACCTAAATAAAGTTTCACCTGAGAGAGAAACAGAAAATATCTCTAAATGG 6234
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrp 220
Qy 6235 GTAATCGTATGCTTCGGGAATCGTATATTATGAGGCTCGGTAAACAACACGAGCTCC 6294
Db 221 ValAsnGlyIleSerTrpGlyLeuValTyrTrpGlySerGlyArgLysLysGlySer 240
Qy 6295 ATTCTAACTATTCGCTCAAAATA---AACACAGCTGGAGCTCCCAATGGCTATAGGACCA 6351
Db 241 ValLeuThrIleArgLeuArgGlieGluThrGlnMetGluProProValAlaIleGlyPro 260
Qy 6352 AATACGGTCTTGACGGGTCAAGACGCCCAACCCCA-----GGACACGAGACCA 6399
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGlnArgProSerProAsnPro 280
Qy 6400 TCCTCT---AACATACTCTCGATCAGACCCCACTGAGTCTAACACACGACTAAATG 6456
Db 281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
Qy 6457 GGGGCAAAATTTTAGCCTCATCCAGGAGCTTTTCAAGCTCTTAAGTCCACGACTCCA 6516
Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
Qy 6517 GAGGCTACCTCTCTTGTGCTATGCTTAGCTTCGGGCCCACTTATGTAAGGAATG 6576
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTrpGluGlyMet 340
Qy 6577 GCTAGAGAGGGAATTCATGTGCAAAAGACATAGACCAATGCCATCGGGATCC 6636
Db 341 AlaArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
Qy 6637 CAAATAAGCTTACCTCTACTGAGGTTTCTGGAAGGCACTGCATAGGAAAGTTCCC 6696
Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380
Qy 6697 CCATCCCAACCAACCTTTGTATACCACTGAAGCCCTTTAATCAAACTCTGAGAGTCAA 6756
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
Qy 6757 TATCTGGTACCTGGTTATGACAGCTGCTGGGCACTGTAATCTGATTAACCCCTTGTT 6816
Db 401 TyrLeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 420
Qy 6817 TCCACCTTGGTGTTTTAAACCAAACTAAAGATTTTTCATTATGTCCTCAAAATGTTCCCGA 6876
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
Qy 6877 GTGTATTACTATCCCGAAAGCAATCTTTGATGAATATGACTACAGAAATCATCGACAA 6936
Db 441 ValTyrTyrProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460
Qy 6937 AAGAGAACCCCATATCTGACACTGCTGTGATGCTGGAGTGGAGTGGCAGAGT 6996
Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyValAlaIleGly 480
Qy 6997 GTAGGAACAGGAACAGCTGCTCGGTACGGGACCAACAGCAGCTAGAAAACAGGACTTGT 7056
Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
Qy 7057 AACCTACATCGAATTCTGAAGAAGATCTCCAAGCCCTAGAAAATCTGTCAGTAACTTG 7116
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
Qy 7117 GAGGAATCCCTAACCTCTTATCTGAAGTAGTCTTACAGATGAGAGAGGGTTAGATTTA 7176
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgGlyLeuAspLeu 540
```

```
Qy 7177 TTATTTCTAAAGAGGAGGATTATGTAGCCTTGAAGAGGAATGCTGTTTTATGTG 7236
Db 541 LeuPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrVal 560
Qy 7237 GATCATTCAGGGGCCATCAGAGACTCCATCAACAAAGCTTAGAGAAAGTTGGAGAACGT 7296
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 580
Qy 7297 CGAAGGGAAGCAAACTACTCAAGGGTGGTGTGAGGGATGGTTCACACAGTCTCTTTGG 7356
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
Qy 7357 TTGGCTACCTTCTTCTTAAACAGGACCTTAAATAGTCTCTCTCTCTCTCTCTCTCTCTCT 7416
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThr 620
Qy 7417 GTTGGGCCATGTATTATTAACTTAATTCCTTCAATTAGAGAACGAATAAGTGCAGTC 7476
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
Qy 7477 CAGATCATGTACTTATAGACAAACAGTACCAA 7506
Db 641 GlnIleMetValLeuArgGlnGlnTyrGln 650
```

Search completed: February 14, 2006, 17:17:01

Job time : 1161.38 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:19:01 ; Search time 7.13096 Seconds
(without alignments)
2992.933 Million cell updates/sec

Title: US-10-723-552-3
Perfect score: 14636
Sequence: 1 GCGTGTGTACGACTGTGGG.....CTGTTTGCATCAAAAAAAA 8132

Scoring table: BLOSUM62
-O=/abss/ABSSWEB.spool/US10723552/runat.14022006.125151.13368/app.query.fasta_1
-DB=Published Applications_AA_New -QFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -HOST=abss02p
-USER=US10723552.@CGN.1.1.17 @runat.14022006.125151.13368 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Searched: 97014 seqs, 1312538 residues

Total number of hits satisfying chosen parameters: 194028

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abss/ABSSWEB.spool/US10723552/runat.14022006.125151.13368/app.query.fasta_1
-DB=Published Applications_AA_New -QFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -HOST=abss02p
-USER=US10723552.@CGN.1.1.17 @runat.14022006.125151.13368 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_New:

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pap:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1735	11.9	1189	US-10-821-234-1209	Sequence 1209, Ap
2	1646.5	11.2	1150	US-10-821-234-1083	Sequence 1083, Ap
3	1358	9.3	537	US-11-129-442-47	Sequence 47, Appl
4	1267.5	8.7	654	US-10-510-947-6	Sequence 6, Appl
5	431	2.9	912	US-11-042-988-12	Sequence 12, Appl
6	413.5	2.8	540	US-10-821-234-1395	Sequence 1395, Ap
7	301	2.1	562	US-10-507-928-10	Sequence 10, Appl
8	301	2.1	562	US-10-507-928-12	Sequence 12, Appl
9	301	2.1	562	US-11-029-465-10	Sequence 10, Appl

10	287.5	2.0	561	7	US-11-029-465-12	Sequence 12, Appl
11	250	1.7	229	7	US-11-014-629-1	Sequence 1, Appl
12	217.5	1.5	210	7	US-11-022-562-221	Sequence 221, App
13	210.5	1.4	422	7	US-11-230-251-26	Sequence 26, Appl
14	199.5	1.4	2760	7	US-11-124-367A-444	Sequence 444, App
15	199.5	1.4	2803	7	US-11-124-367A-442	Sequence 442, App
16	199.5	1.4	2803	7	US-11-124-367A-445	Sequence 445, App
17	199.5	1.4	2984	7	US-11-124-367A-443	Sequence 443, App
18	199.5	1.4	3027	7	US-11-124-367A-441	Sequence 441, App
19	181.5	1.2	2108	6	US-10-995-561-989	Sequence 989, App
20	181	1.2	1823	6	US-10-995-561-988	Sequence 988, App
21	181	1.2	2102	6	US-10-995-561-990	Sequence 990, App
22	181	1.2	2157	6	US-10-995-561-991	Sequence 991, App
23	174.5	1.2	188	7	US-11-234-786-592	Sequence 592, App
24	167	1.1	14130	7	US-11-175-689-9	Sequence 9, Appl
25	163.5	1.1	416	7	US-11-226-657-136	Sequence 136, App
26	158.5	1.1	16990	7	US-11-175-689-7	Sequence 7, Appl
27	157.5	1.1	1558	6	US-10-329-258-14	Sequence 62, Appl
28	155.5	1.1	2766	6	US-10-877-346-62	Sequence 19, Appl
29	152	1.0	1730	7	US-11-182-016-19	Sequence 8, Appl
30	151.5	1.0	1516	6	US-10-220-824-8	Sequence 1, Appl
31	151.5	1.0	1560	7	US-11-059-982-1	Sequence 50, Appl
32	150.5	1.0	3073	7	US-11-143-980-50	Sequence 20, Appl
33	150	1.0	1717	7	US-11-182-016-20	Sequence 20, Appl
34	148	1.0	1717	7	US-11-182-016-20	Sequence 1143, Ap
35	147.5	1.0	1388	6	US-10-821-234-1143	Sequence 2, Appl
36	147	1.0	1618	6	US-10-984-645-2	Sequence 1155, Ap
37	147	1.0	4419	6	US-10-821-234-1155	Sequence 155, App
38	144.5	1.0	400	7	US-11-186-284-155	Sequence 27, Appl
39	144	1.0	2811	6	US-10-877-346-27	Sequence 1916, Ap
40	143.5	1.0	1076	6	US-10-467-657-7916	Sequence 19, Appl
41	143	1.0	1730	7	US-11-182-016-19	Sequence 1050, Ap
42	142.5	1.0	1410	6	US-10-821-234-1050	Sequence 25, Appl
43	142.5	1.0	2814	6	US-10-877-346-25	Sequence 23, Appl
44	142	1.0	1742	7	US-11-182-016-23	Sequence 56, Appl
45	142	1.0	2011	7	US-11-080-991-56	

ALIGNMENTS

RESULT 1
US-10-821-234-1209
; Sequence 1209, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1209
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1189)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-821-234-1209

Alignment Scores:
Pred. No.: 1.17e-113 Length: 1189
Score: 1735.00 Matches: 453
Percent Similarity: 50.9% Conservative: 192
Best Local Similarity: 35.8% Mismatches: 462
Query Match: 11.9% Indels: 160

DB:	6	Gaps:	34
US-10-723-552-3	(1-8132) x US-10-821-234-1209 (1-1189)		
QY	2121	GGACCAAGGATCCTAGCTCTAGAGAAGATAAAGATTAGGGAGACGGGTTCCGACCCC	2180
		:::	
Db	11	GlyProArgLeu-----	pro 15
QY	2181	CTCCCC-----GAGCCAGGGTAACTTTGAAGGTGGAGGGCAACCA	2222
		:::	:::
Db	16	AlaProThrAlaIleThrAlaSerGluProArgValThrLeuLeuValAlaGlyLeuLeu	35
QY	2223	GTTGAGTTCCTGGTTGATACCGGAGCGAAACATTACGTGCTACTACAGCCATTAGGAAAA	2282
		::: :::	:::
Db	36	IleSerPheLeuIleAspSerGlyAlaThrTyrSerAlaLeuProGluPheSerGlyPro	55
QY	2283	CTAAAGATATAAAATCTGGTGATGGGTCCACAGGGCAACAACAGTATCCATGGACT	2342
		::: :::	
Db	56	ThrHisProSerGlnValSerValValGlyValAspGlyLeuIleSerCysProLeuAla	75
QY	2343	ACCCGA--AGAACAGTTGACTTCGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCTATA	2399
		:::	:::
Db	76	ThrArgProLeuThrCysSerLeuPheAspAlaIlePheSerHisSerPheLeuIleMet	95
QY	2400	CCTGAGTCCCAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAAATT	2459
		:::	
Db	96	ProHisCysProThrProIleLeuGly**AspLeuLeuAlaIleValIleAlaSerIle	115
QY	2460	TCCTTTTGAA-----CAAGGGAAACAGAAAGTGCTCGCAATATACAAACCT	2504
		::: :::	
Db	116	ThrPheSerCysProLeuGlnProGluSerProProAlaProProAlaSerProAlaPro	135
QY	2505	ATCACTGTGTGACCCCTCAATTAGATGACGAATATGACTATACTCTCCCTAGTAGAAG	2564
		:::	
Db	136	AspProPhe-----ProGlnTyrProLeuProAlaSerLeuIleAsn	149
QY	2565	CCTGATCAAAATATACAATTCTGTTGGACAGTTTCCCAAGCTGGGAGAAACCGCA	2624
		:::	
Db	150	Pro-----	150
QY	2625	GGGATGGGTTTGGCAAGCAAGTTCCCCCAACAAGTTATTCACTGAAGCGCAGTCCACA	2684
		:::	:::
Db	151	-----SerValAlaAlaHisHisAspProIleArgIleGlnLeuLysAspSerSerLys	168
QY	2685	CCAGTGTCACTCAGACAGTACCCCTCTGAGTAAAGACTCAAGAAGAAATTCGCCCGCAT	2744
		:::	
Db	169	PheProSerValProGlnTyrProIleSerLeuThrHisGlnLysGlyLeuGlnProIle	188
QY	2745	GTCCAAAGATTAAATCAACAGGGCATCTAGTTCCTCTCCAATCTCCCTCGAATACCTCCC	2804
		:::	:::
Db	189	ValAsnLysLeuCysSerCysSerLeuLeuArgProThrHisSerProTyrAsnThrPro	208
QY	2805	CTGCTACCGGTTAGAAAGCCCTGGGACTAANGACTATCGACAGTACAGGACTTGAGAGAG	2864
		::: :::	
Db	209	IleLeuProValLysLys--SerAspGlySerTyrArgLeuValHisAsnLeuGlnAla	227
QY	2865	GTCAATAAACGGGTGCAGGATATACACCCACAGTCCCGAACCTTATAACCTCTTGTGT	2924
		::: :::	
Db	228	IleSerGlnAlaValLeuProIleHisProIleValHisAsnProTyrProLeuLeuSer	247
QY	2925	GCTCTCCCAACCCCAACGAGCTGTGTATACAGTATTGGACTTTAAAGGATGCCTTCTTCGC	2984
		::: :::	
Db	248	LeuValAlaThrAsnThrThrLeuTyrThrAlaIleAspLeuLysAspAlaPheThr	267
QY	2985	CTGAGATTACACCCCACTAGCCAACCACTTTTTCCTTCGAATCGAGAGATCCAGGTACG	3044
		::: :::	
Db	268	IleSerLeuHisProAspSerGlnAsnLeuPheAlaPheThrTrpThrAspProAspThr	287
QY	3045	GGAAGAACCGGCGACTCACCTGGACCCGACTGCCCAAGGGTTCAAGACTCCCCGACC	3104
		::: :::	
Db	288	LeuGlnSerGlnGlnLeuThrTrpThrValLeuProGlnGlyPheArgSerProSer	307
QY	3105	ATCTTTGACGAAGCCCTACACAGAGACTGGCCCAACTTCAGGATCCCAACCCCTCAGGTG	3164

Db 655 ProLeuThrLys***LeuLysProGlySerSerSerGlnLysAlaGluLeuLeuAlaLeu 674
 QY 4212 ACACAAGCTTTGGCGTGGCGAAGGAAATCCATAACATTTATATACGACAGCAGGTAT 4271
 Db 675 ThrAlaLeuAenLeuSerLysGlyLysArgValAsnIleTyrThrAspSerLysTyr 694
 QY 4272 GCCTTTCCGACTGACACAGTACATGGCGGCATCTATAACAAGAGGGGGTTCCTACCTCA 4331
 Db 695 AlaTyrHisIleProArgSerHisAlaAlaIleTyrGlnLysArgGlyLeuLeuThrAla 714
 QY 4332 GCAGGAGGGAATAAAGACAAAGAGGAATCTTAAGCCTATTAGAAGCGGTACATTTA 4391
 Db 715 LysGlyThrProIleThrAsnGlyHisLeuIleTyr***LeuLeuGlnAlaThrHisLeu 734
 QY 4392 CCAAAAAGGCTAGCTATTATACACTGCTCGACATCAGAAAGCTAAGATCTCATATCC 4451
 Db 735 ProAlaLysAlaGlyValIleHisCys**GlyHisArgThrGlySerAspGluIleSer 754
 QY 4452 AGAGGAACCATGATGGCTGACCGGGTTCACAGCAGGCAGCCCGGGTGTAACTTCTG 4511
 Db 755 LysGlyAsnArgLysThrAspGluAlaAlaLysGlnAspSerLeuSerProLeuProAla 774
 QY 4512 CCTATATAGAAATGCCAAAGCCCAAGACCCAGACAGTACACCTTAGAAGACTGG 4571
 Db 775 ProIleLeuLeuValThrProAlaValProProArgTyrProProThr----- 790
 QY 4572 CAAGAGATAAAAAGATAGACCAAGTCTCTGAGACTCCGGAAGGGAGCTCTATACCTCA 4631
 Db 791 ---GluLysSerSerLeuLeuGlnGlnGlyAlaSerLeuGlnGlyAspTyrIleLys 809
 QY 4632 GATGGGAGAAATCTGCCCCACAAAGAGGGTTAGAAATATGCCAACAGATACATCGT 4691
 Db 810 AsnGlnLysProValLeuProGlnGlnSerLysGluIleLeuThrProLeuHisGln 829
 QY 4692 CTAACCCACTAGGAACCTAACACCTGCAGCAGTGTGTGAGA-----ACATCCCCT 4742
 Db 830 ProPheHisIleSerAlaCysProLeuTyrLeuLeuLeuArgProCysPheSerSerPro 849
 QY 4743 TATCATGTTCTGAGGCTTACAGGAGTGGCTGACTCGGTGTCAACATTTGTGTGCCCTGC 4802
 Db 850 HisLeuPheIleSerLeu-----LysAspIleThrSerAsnCysArgIleCys 865
 QY 4803 CAGCTGGTTAATGCTAATCTTCCAGAAATGCCT-----CCAGGAAGAGACTAAGG 4853
 Db 866 SerValThrSerSerGlnGlyAlaLeuCysProLeuLeuLeuIleLeuThrTyrGlnLeuArg 885
 QY 4854 GGAAGCCACCCAGCGCTCACTGGAGTGGACTCACTGAGGTAAAGCCGGTAAATAC 4913
 Db 886 GlyThrLeuProGlyGluHisTyrGlnValAsnPheThrHisMetProProValLys--- 904
 QY 4914 GGAACAAATACCTATTGGTTTGTAGACACCTTTTCAGATGGGTAGAGGCTTATCCT 4973
 Db 905 LysSerLysTyrLeuLeuThrLeuValAspThrPheSerGly**ValGluAlaPhePro 924
 QY 4974 ACTAAGAAAGAGACTTCAACCGTGGTGTGCTTAAAAAATACTGGAAGAAATTTTCAAGA 5033
 Db 925 ThrProSerGlyLysAlaAlaGluValSerGlnIleLeuValThrGluIleLeuProArg 944
 QY 5034 TTTGGATACCTAAGGTAAATAGGTGAGCAATGTGCTCAGCTTTTGTGCCAGGTAACT 5093
 Db 945 PheGlyLeuProGlySerIleGlnSerAspAsnSerProSerPheIleSerGlnIleThr 964
 QY 5094 CAGGAGACTGCCAAGATATTGGGATTTGATTGAAACTGATGTGCATACAGACCCCAA 5153
 Db 965 GlnGlnValSerGlnSerLeuGlyIleGlnTyrArgLeuHisIleProCysTyrProGln 984
 QY 5154 AGCTCAGGACAGGTAGAGGATGAATAGAACCATTAAGAGACCTTACTAAATGACC 5213
 Db 985 ThrSerGlyLysValGluArgAlaAsnGlyIleLeuLysAlaGlnLeuThrLysLeuThr 1004
 QY 5214 GCGGAGACTGGCGTTAATGATTGGATAGTCTCTGCTCCCTTTGTGCTTTTGGGTAGG 5273
 Db 1005 LeuGluVal---GlnLysProTyrThrSerLeuLeuProIleAlaLeuGluSerIleArg 1023

QY 5274 AACACCCCT---GGACAGTTTGGGCTGACCCCTATGAATTTACTCTACGGGGACCC--- 5327
 Db 1024 AlaSerProLysAlaProSerPheLeuSerProPheGluLeuIleTyrGlyArgProPhe 1043
 QY 5328 -----CCCCCATTTGGTAGAAATTTGCTTCTGTATACATAGTGTGACGTGCTG 5372
 Db 1044 LeuLeuGlnAsnArgProPro-----SerAsnSerGlnLeuGlyGluTyrLeu 1059
 QY 5373 CTTTCCAGCCTTTGTTCTCTAGCTCAAGGCACCTTGAGTGGGTGAGACACAGCGCTGG 5432
 Db 1060 ProThrValSerLeuMetSerTyrLeu-----LeuCysGlnGlnAlaAsp 1074
 QY 5433 AGCAAACTCCGGAGCGCTACTCTCAGGAGGAGGAGACTTTCAG-----ATC 5477
 Db 1075 GlnAlaLeuProLysProHisGluGlyValSerAsnProLys**ThrCysSerProIle 1094
 QY 5478 CCACATCGTTTCCAACTGGGAGATTCA-----GTCTACGTTAGACCCACCGTGCA 5528
 Db 1095 ProLys-----AspSerLeuSerArgValThrLeuGlnAsnHisArgGly 1109
 QY 5529 GGAACCTCGAGACTCGGTGGAGGCGCCCTTATCTGCTACTTTTG----- 5573
 Db 1110 LeuAspLeuLeuThrAlaGluLysGlyLeuCysIlePheLeuGluGluCysCys 1129
 QY 5574 -----ACCACACCAACGGCTGTG-----AAAGTC 5597
 Db 1130 PheTyrThrAsnGlnSerGlyLeuValGlnAspAlaAlaGlyArgIleAsnGluLysAla 1149
 QY 5598 GAAGGAATCTCCACCTGGATCCATGTCATCCACGTTAAACCGCGGCACCTCCCGATTCG 5657
 Db 1150 SerGlyArgValGlnTrpLeu-----ThrProValIleProAsp 1162
 QY 5658 GGTGTGAAAGCCGAAAG-----ACTGAAAAT 5684
 Db 1163 LeuTrpGluAlaGluAlaGlyGlySerArgGlyGlnGluIleGluThrIleLeuAlaAsn 1182
 QY 5685 CCCCTTAAGCTTCGCTCCAT 5705
 Db 1183 ThrValLysProArgLeuTyr 1189

RESULT 2

US-10-821-234-1083
 ; Sequence 1083, Application US/10821234
 ; Publication No. US20050255114A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Labat, Ivan
 ; APPLICANT: Stache-Crain, Birgit
 ; APPLICANT: Andarmani, Susan
 ; APPLICANT: Tang, Y. Tom
 ; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
 ; FILE REFERENCE: 821A
 ; CURRENT APPLICATION NUMBER: US/10/821,234
 ; CURRENT FILING DATE: 2004-04-07
 ; PRIOR APPLICATION NUMBER: US 60/462,047
 ; PRIOR FILING DATE: 2003-04-07
 ; NUMBER OF SEQ ID NOS: 1704
 ; SOFTWARE: pt_seq_genes Version 1.0
 ; SEQ ID NO 1083
 ; LENGTH: 1150
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(1150)
 ; OTHER INFORMATION: xaa = any amino acid or nothing
 US-10-821-234-1083

Alignment Scores:
 Pred. No.: 1,77e-107 Length: 1150
 Score: 1646.50 Matches: 438
 Percent Similarity: 51.9% Conservative: 186
 Best Local Similarity: 36.4% Mismatches: 470

Query Match:	11.2%	Indels:	109
DB:	6	Gaps:	37
US-10-723-552-3 (1-8132) x US-10-821-234-1083 (1-1150)			
Qy	2226	GAGTCTCTGGTTGATACCGGAGCGAAACATTCAGTGTCTACTACAGCCATTAGGAAAACTA	2285
Db	4	GlnPheLeuValAspSerArgAlaAaCysSerValLeuThrGlnProIleCysLeuLeu	23
Qy	2286	AAAGATAAAATCCTGGGTGATGGTGCCACAGGCAACACAGTATCCATGGCATACC	2345
Db	24	SerAsnCysAsnCysArgValMetGlyIle-----AspArgCysProLysValArg	40
Qy	2346	CGAAGACAGTTGACTGGAGTGGGAGCGGTATACCCAC-----TCGTTTCTG	2393
Db	41	LysPheThrPheProLeuAlaCysGluAlaMetSerArgLeuPheSerProHisPheLeu	60
Qy	2394	GTCACTACTGAGTCCCGACAGCCCTCTTAGGTAGAGACTTATTAGCAAGATGGGAGCA	2453
Db	61	CysProGluCysProThrProLeuLeuGlyArgAspLeuLeuSerLysMetGlyAla	80
Qy	2454	CAATTTCTTTGAACAAGGAAACACAGAAGTGTCTGCAAT---AACAAACCTATCACT	2510
Db	81	ThrIleSerLeuGluGluAspArgLeuGlnValGluAlaGluProGluGlnGlyIleHis	100
Qy	2511	GTGTTGACCTCCAAATAGATGACGAATATCGACTATACTCTCCCTAGTAAGCCCTGAT	2570
Db	101	LeuLeuAlaLeuLeuAsnGlyGlnGluLeuGlu-----Thr	112
Qy	2571	CAAAATATA---CAATCTCTGGTTCGAAACAGTTTCCCAAGCCTGGCGAGAAACCGCAGGG	2627
Db	113	GlnAsnIleProLysGluSerLysAspHisIleThrProSerLeuSerAspThrSerVal	132
Qy	2628	ATGGTTTGGCAAGCAAGTTTCCCCACAAAGTTATTCAACTGAAGGCCAGTGCCACACCA	2687
Db	133	LeuGlyGlnAlaAsnLysValProProValLysThrAspLeuLysProGlyMetGlyTyr	152
Qy	2688	GTGTCACTCAGACAGTACCCCTTGAGTAAGAAGCTCAAGAGGAATTCGCGCGCATGTC	2747
Db	153	ProTrpArgLysSerTyrLeuLeuLysProValAlaLeuAsnGlyValGlnProLeuLeu	172
Qy	2748	CAAAAGATTAAATCCAAAGGGCATCTAGTTCCTGCTCAATCTCCCTCGAATACTCCCTCG	2807
Db	173	HisLysPheLeu---GlnGly---LeuArgProCysGlnSerProGlyIleAsnThrSerVal	191
Qy	2808	CTACCGGTTAGAAAGCTGGGACTTAATGACTATATCGACCATCGACGATTCAGAGAGGTC	2867
Db	192	LeuProValLysLysPro---AsnGlyGluTyrGlnPheVal***AspLeuLysGlyVal	210
Qy	2868	AATAAACGGGTGAGGATATACACCCACAGCTCCGCAACCTTATAACCTCTTGTGTGCT	2927
Db	211	AsnGluAlaValIleProIleHisProThr---ProGlyPro---LysLeuValCysAla	228
Qy	2928	---CTCCCAACCCCAACGGAGCTGGTATACAGTATTGGACTTAAAGGATGCTCTTCTGCG	2984
Db	229	GlnIleProGlyAspAlaGlnPhePheThrLeuLeuHisPheLysAspAlaPhePheCys	248
Qy	2985	CTGAGATTACACCCCACTAGCCAAACCACTTTTTCCTTCGCAATGGAGAGATCCAGGTACG	3044
Db	249	ValAlaLeuSerThrGlnAsnProCysIlePheCysPheGluTyrArgAsnProAspThr	268
Qy	3045	GGAAGAACCGGGAGCTCACTGACCCGAGCTGCCCAAGGTTCAAGAACTCCCGACC	3104
Db	269	GlnLysAlaThr***TyrTyrTrpThrValLeuProGlnGlyPheTrpGluGlyProHis	288
Qy	3105	ATCTTTGACGAAGCCCTACACAGAGACTCGGCCAACTTCAGGATCCAAACCCCTCAGGTG	3164
Db	289	LeuPheGlyAsnAla-----GlySer***AlaTyrArgMetGly-----	301
Qy	3165	ACCTCTCTCAGTACGTGGATGACTGCTTCTGGGGAGGCCCAACAGGACTGCTTA	3224
Db	302	LeuLeuLeuGlnTyrValAspAspLeuPheIleValSerGlnThrTrpGlnAspSerAsp	321

Qy	3225	GAAGGCACGAAGCACCTACTCTGCTGGAATTTGCTGCACTAGGCTACAGAGCCTCTCTAAG	3284
Db	322	PheAsnIleIleLysThrLeuAsnPheLeuAlaGluArg***TyrLysAlaAlaProSer	341
Qy	3285	AAGCCCCAGATTTTCAGGAGAGAGGTAAACATCTTGGGGTACAGTTTTCGGGACGGGCAG	3344
Db	342	LysAlaGlnIleSerLeuGlnLysPheArgTyrLeuGlyPheIleLeuThrProGlyAla	361
Qy	3345	CGATGGCTGACGAGGACGGAAGAAAACTGTAGTCAGATACCGGCCCAACACACAGCC	3404
Db	362	ArgThrLeuValAspGlyLeuGluLysAlaIleThrSerLeuLeuValProGlnThrLys	381
Qy	3405	AAACAATAGACAGTTTTGGGGACAGCTGGATTTTGACAGCTGTGGATCCCGGGGTTT	3464
Db	382	ArgGlnProGlnGlyPheLeuGlyMetValGlyPheCysSerIleTrpIleProAsnTyr	401
Qy	3465	CGACCTTAGCAGCCCTCTACTACCGCTAAACCAAA-----GAAAAAGGGAATTCCTCC	3518
Db	402	GlyLeuThrAlaLysProLeuCysGluThrGlnLysGlyGluArgGluProLeuTyr	421
Qy	3519	TGGGCTCTGAGCACACAGAGGCATTTGATGCTATCAAAAAGGCCCTGTGAGCGGCCT	3578
Db	422	TrpGluLysLysCysGlnLeuSerPheGluProLeuLysThrGluLeuGlyGlnAlaLeu	441
Qy	3579	GCTCTGCGCTCCCTCAGCTAAACCCCTTTACCTTTATGTTGATGAGCGCTAAGGGA	3638
Db	442	ValLeuGlyLeuProAsp---GluLysProLeuThrLeuTyrValArgGluArgLeuGly	460
Qy	3639	GTAGCCCGGGAGCTTTTAACCCAAACCTTAGCACCATGGAGAGAGCTGTGCGCTACCTG	3698
Db	461	IleAlaLeuGlyValLeuThrGlnArgLeuGlyProValGlnLysProValAlaHisPhe	480
Qy	3699	TCAAAGAGCTCCATCCTCTAGCCAGTGGTTGGCCCATATGCTGAAGGCTATCGCAGCT	3758
Db	481	SerLysGlnThrAsnTyrProValAlaGlnLysArgProGlyCysTrpArgAlaValAlaAla	500
Qy	3759	GTGCGCATCTGTCAAGGACGCTGACAAATTCACCTTTGGGACAGATAATACTGTATAA	3818
Db	501	ThrAlaLeuLeuPheSerGluAlaSerLysLeuThrGlnGlyGlnTyrThrGluAlaMet	520
Qy	3819	GCCTCCCATCTGGAGAACATCGTTTGGGACGCCCCCAGACCGATGGATGACCAACGCC	3878
Db	521	ThrProHisGlnValGlnThrGlyLeuGluValLysGlyHisIleStrpLeuThrAlaGly	540
Qy	3879	CGATGACCCATCATCAAGCCCTGCTTCTC---ACAGAGAGGCTCAGTTCGCTCCACCA	3935
Db	541	ArgLeuThrAsnTyrGlnAlaLeuLeuLeuAspThrLeuAspIleThrLeuArgValCys	560
Qy	3936	GCCGCTCTCAACCTCGCCACTCTTCTGCTGAAGAGACTCATCAACCCAGTCTCATGAT	3995
Db	561	GlnThrLeuAsnThrAlaThrSerLeuProThrThrAspGluAspSerGlnHisGln	580
Qy	3996	TGCCATCAACTATTATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTG	4055
Db	581	CysThrGly---ProThrLysGlnThrHisSerArgProSerLeu***AspLysProLeu	599
Qy	4056	ACTGGAGAGTGTCTACCTGGTTCACTGACGGAAGCAGCTATGTGTGGAGAGGTAAAGAG	4115
Db	600	GluAsnAlaLysThrGlu***TyrProAspGlySerSerPhe---ValGluGlyThrGln	618
Qy	4116	ATGCTGGGGCGCGGTGGGACGGGACCCCGCACGATCTGGGCCAGCAGCTGCCGGAA	4175
Db	619	LysAlaGlyHisAlaValValSerLeuLysGluThrThrGluThrLysAlaLeuProPro	638
Qy	4176	GGAATTTACACAAAAGGCTGAGCTCATGCGCCCTCACGCAAGCTTTGCGGCTGGCGAA	4235
Db	639	GlnThrSerAlaGlnLysAlaAlaLeuGluAlaLeuThrGluProHisAsnTrpGluLys	658
Qy	4236	GGAATATCAATAACATTTATACGACAGCAGGTATGCCCTTGGACTGCACACGTACAT	4295
Db	659	GlyArgArgValAsnLeuTyrThrAsnSerLysTyrArgPheLeuLeuLeuCysAlaHis	678
Qy	4296	GGGCGCATCTATAAAACAAAGGGGGTGTCTTACCTCAGCAGGAGGGGAAATAAAGACAA	4355

```
Db 679 AlaAlaIleTrpLysGluLysGlyMetLeuThrAspLysAsnSerProGlyLysArgLys 698
QY 4356 GAGGAAATCTTACAGCCATTAGAGCCGTACATTTACCAAAAGAGCTAGCTATTATACAC 4415
Db 699 AspProIleGlnLeuLeuLeuGluAlaGlnLeuProCysGlnValAlaValIleHis 718
QY 4416 TGTCCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATGGCTGACCG 4475
Db 719 CysLysGlyHisGlnArgAspTrpSerLeuIleSerGlnGlyAsn-----Arg 734
QY 4476 GTTGCCCAAGCAGGCA-----GCCAGGGTGTAACTTCTGCCTATAATAAGAATG 4526
Db 735 AlaThrLysGlnAlaThrTrpLeu**GlusSerAsnValLeuAlaLeuMetThrAsp 754
QY 4527 CCCAAAGCCCAAGACCCAGACAGCAGTACACCCCTAGAACCTGGCAAGAGATAAAAG 4586
Db 755 ProThrAsnLeuProAsnHisLeuGlnTySerGlnGlnGlu---GlnGlu**AlaLys 773
QY 4587 ATAGACAGTTCTCTGAGACTCCGGAAGGACCTGTATACCTCAGATGGGAAGAAATC 4646
Db 774 GlyGlnGlyTyTrpGlnGluAsnSerAlaGly-----TrpPheThrAspGlyGluSerPhe 791
QY 4647 CTGCCCCACAAGAGGGTATAGATATGTCCACAGATACATCGTCTAACCCACCTAGGA 4706
Db 792 IleProGlyThrAspGlnLeuLysValThrLysHisLeuHisAspAlaThrHis----- 809
QY 4707 ACTAAACACCTGCAGCAGTGTGTGCAGAACATCCCTTATCATGTCTTGAGGCTACCAGGA 4766
Db 810 -----**TrpArgAspValLeuTrpAsnSerValHisLysIlePheThrArgLysLys 827
QY 4767 GTGGCTGACTCGTGTGTCACAAATGTGTGCTCCCTGCAGCTGGTTAATGTAATCTCTTC 4826
Db 828 LeuArgLysSerValLysGlnValThrPheValSerGlnLeuCysSerArgAsnPro 847
QY 4827 AGA-----ATGCTCCAGGGAAG-----AGACTAAGGGGAAGCCACCCA 4865
Db 848 HisThrHisProIleProSer**LeuLysLeuValGlnHisArgGlyThrTyLeu 867
QY 4866 GCGCGCTCAGTGGGAAGTGGACTTCACTGAGTAAAGCGCGTAAATACGGAACAATAC 4925
Db 868 GlyLysAspTrpGlnValValLeuThrGlnMetThrPro---AsnLeuGlyTyTrpLysTy 886
QY 4926 CTATTGGTTTTGTAGACACTTTTCAAGATGGGTAGAGCTTATCTACTAAGAAGAG 4985
Db 887 LeuLeuValPheValAspThrPheThrGlyArgValLysGlyPheProThrCysThrGlu 906
QY 4986 ACTTCAACCTGGTGGCTTAAATAATCTCGAAGAAATTTTCAAGATTGGAATACCT 5045
Db 907 LysAlaValGluValCysLysProSerLeuLysGluValIleSerGlnPheGlyLeuPro 926
QY 5046 AAGGTAATAGGTGACAGACATGTGTCCAGCTTTTGTGCCAGGTAAAGTCAAGGAGCTGGCC 5105
Db 927 LysSerProGlnSerGlyAsnArgLysLeuSerPheMetGlyLysIleThrGlnSerLeuSer 946
QY 5106 AAGATATTGGGGATTGATTGGAAACTGCATTTGTGCATACAGACCCCAAGCTCAGGACAG 5165
Db 947 ThrThrLeuGlyIleAspTyTrpGlnLeu-----AlaProGlnSerSerGlyLys 962
QY 5166 GTAGAGAGGATGAATAGAACCATTAAGAGACCTTACTAAATTTGACCGCGGAGACTGGC 5225
Db 963 Val---LysMetAsnHisThrLeuGluThrThrLeuAlaLysLeuPheGlnGluMetHis 981
QY 5226 GTTAATGATTGGATAGCTCTCTCGCCCTTTTGTGCTTTTATAGGCTTAGGAACCCCTGGA 5285
Db 982 ---GlusSerTrpValLysMetLeu---LeuSerLeuLeuArgValArgAlaProSer 999
QY 5286 ---CAGTTTGGGCTGACCCCTATGAATTACTCTACGGGGGACCCCCCATTTGGTAGAA 5342
Db 1000 CysSerLeuArgLeuSerSerGlnMetIleTyTrpGlnArgPro----- 1014
QY 5343 ATTGCTTCTGTACATAGTGTGTGCTGCTTCTCCAGGCTTTTGTCTCTAGGCTCAAG 5402
```

```
Db 1015 -----PheLeuThrThrAspLeuLeuLeuGlyGluLeuHisLysLysLeu--- 1030
QY 5403 GCACCTTGAGTGGGTGAGACAAACGAGCGTGAGGCAACTCCGGGAGGCTTACTCAGGAGGA 5462
Db 1031 -----GlnTyArgPheArgThrAspSerLys 1039
QY 5463 GAGACTTGCAGATC-----CCACATCGT----- 5486
Db 1040 GlyAspPheArgIleTrpLysGlnAsnProAlaSerProHis**GlyTrpGluAsnPro 1059
QY 5487 TTCCAAGTG-----GGAGATTACAGTCTAGACGCCACCGTCGAGGAAACCTCGAG 5540
Db 1060 PheGlnValAsnProGlyAspGlnValLeuLeuLysAla***ArgAlaGlySerProGlu 1079
QY 5541 -----ACTCGGTGAAAGGCCCTTATCTCGTACTTTTGACCAACCAACCGCT 5588
Db 1080 AspHisProLeuLeuLysTrpGluGlyProCysTrpValIleLeuThrThrProThrAla 1099
QY 5589 GTCAAAAGTCGAAGAATCTCCACCTGGATCCATGTCATCCACGTTAAACCGGCGCCACCT 5648
Db 1100 AlaAsnSerGlnGlyIleThrSerTrpValHisLeuSerArgSerGluMetLeuSerPro 1119
QY 5649 -----CCGATTCCGGGTGAAAGCCGAAAGACTGAAATCCCTT 5690
Db 1120 LysCysPheGlnThrArgProAspGlyProPheTySer***LysProValGluAspLeu 1139
QY 5691 AAGCTTCGC 5699
Db 1140 LysPheArg 1142

RESULT 3
US-11-129-442-47
; Sequence 47, Application US/11129442
; Publication No. US20060002951A1
; GENERAL INFORMATION:
; APPLICANT: Kleiman, Lawrence
; APPLICANT: CEN, Shan
; APPLICANT: GUO, Fei
; TITLE OF INVENTION: Inhibition of the tRNAlys3-Primed Initiation of Reverse
; FILE REFERENCE: 11168.257
; CURRENT APPLICATION NUMBER: US/11/129,442
; CURRENT FILING DATE: 2005-05-16
; PRIOR APPLICATION NUMBER: CA 2,467,312
; PRIOR FILING DATE: 2004-05-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Murine leukemia virus
US-11-129-442-47

Alignment Scores:
Pred. No.: 2,05e-87 Length: 537
Score: 1358.00 Matches: 277
Percent Similarity: 62.7% Conservative: 76
Best Local Similarity: 49.2% Mismatches: 142
Query Match: 9.3% Indels: 68
DB: 12 Gaps: 12

US-10-723-552-3 (1-8132) x US-11-129-442-47 (1-537)
QY 585 ATGGGACAGACCGTGGACCGCCCTCTAGTTTGTACTCTCGACCATTTGGACTGAAGTTAA 644
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGluHisTrpGlyAspValGln 20
QY 645 TCCAGGCGCTCAATAATTGTTCAGTTTAAAGAGGACCTTGGCAGACTTTCTCTGTGTC 704
Db 21 ArgIleAlaSerAsnGlnSerValAspValLysArgArgTrpValThrPheCysSer 40
QY 705 TCTGAATGCCGACATTCGATGTGGATCGCCATCATCAGAGGGACCTTTAATCTAGATT 764
```

```
Db 41 AlaGluTrpProThrPheGlyValGlyTrpProGlnAspGlyThrPheAsnLeuAspIle 60
Qy 765 ATCTGGCTGTTAAGCAGATTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 824
Db 61 IleLeuGlnValIleValPheSerProGlyProHisGlyHisProAspGlnVal 80
Qy 825 CCCTATATCTTACGTGGCAAGATTTCGGCAGAGATCTCCGCCATCGGTTAAACCATGG 884
Db 81 ProTyrIleValThrTrpGluAlaIleAlaTyrGluProProTrpValIleValProPhe 100
Qy 885 CTGAATAAGCCACAAAGCCAGGTCTCCCGA-----ATTCTGGCTCTTGGAGAGAAAAC 938
Db 101 ValSerProIleValSerProSerProThrGlyProIleLeuProSerGlyProSerThr 120
Qy 939 AAACACTCGGCTGMAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATGTAGGACCA 998
Db 121 GlnProProArgSerAlaLeu---TyrProAlaLeuThrProSerIleLeysProArg 139
Qy 999 CCGCTTGGCCGGAACCCCAATCT-----GTT 1025
Db 140 ProSerIleProGlnValLeuSerAspAsnGlyGlyProLeuIleAspLeuLeuThrGlu 159
Qy 1026 CCCCACCCCTTATCTGGCACAGGTGCGCGAGGGACCTTTGCCCCCTCTCTGGA--- 1082
Db 160 AspProProTyr-----GlyGluGlnGlyProSerSerSerAspGlyAsp 175
Qy 1083 -----GCTCCGCGGTGGAGGACCTGTGTGCA 1109
Db 176 GlyAspArgGluGluAlaThrSerThrSerGluIleProAlaProSerProMetValSer 195
Qy 1110 GGGACTCGGACCGGCGGCGCCACCCCGGAGCGGACAGACAGATCGGCACATTACCG 1169
Db 196 ArgLeuArgGlyIleArgAspProProAlaAlaAspSerThrThrSerArgValPhePro 215
Qy 1170 CTGGCCAGCTACGCCCTCCACACCGGGGGCCAAATTGACGCCCTCCAGTATTGGCCC 1229
Db 216 LeuArgLeuGlyGly-----AsnGlyGln-----LeuGlnTyrTrpPro 228
Qy 1230 TTTTCTCTCGAGATCTCTATAATTGAAACTAACCATCCCTTTCTCGGAGATGCC 1289
Db 229 PheSerSerAspLeuTyrAsnTrpLysAsnAsnAsnProSerPheSerGluAspPro 248
Qy 1290 CAAGCCTCTACGGGTGTGGAGTCTTATGTTCTCTCACAGCCTACTTGGGATGAT 1349
Db 249 GlyIleLeuThrAlaLeuIleGluSerValLeuThrThrHisGlnProThrTrpAspAsp 268
Qy 1350 TGTCAACAGCTGTCTGACAGACTCTTCAACACCGGAGGAGAGAGAATTTCTATTAGAG 1409
Db 269 CysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluLysGlnArgValLeuLeuGlu 288
Qy 1410 GCTAGAAAAAATGTTCTGGGCGCGACGGGCGACCCACGCGGTGCAAAATGAGATTGAC 1469
Db 289 AlaArgLysAlaValArgLysAsnAspGlyArgProThrGlnLeuProAsnGluValAsn 308
Qy 1470 ATGGGATTTCCCTTAACCTGCCCGCGTGGAGTCTACACACCGCTGAGGTAGGAGAGC 1529
Db 309 SerAlaPheProLeuGluArgProAspTrpAspTyrThrProGluGlyArgAsnHis 328
Qy 1530 TTGAAAATCTATCCAGGCTCTGGCGGGTCTCGCGGCGCTCAAGACGCGCCACT 1589
Db 329 LeuValLeuTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGlyArgSerProThr 348
Qy 1590 AATTGGCTAAGTAPAGAAAGTATGATCAGGACCGCAATGAACCCCTCTGTTTTCTT 1649
Db 349 AsnLeuAlaLysValIleThrGlnGlyProAsnGluSerProSerAlaPheLeu 368
Qy 1650 GAGGGCTCTTGGAGCCTTCAGCGGTGACCCCTTTGATCCCACTCAGAGCCCAA 1709
Db 369 GluArgLeuLysGluAlaTyrArgTyrThrProTyrAspProGluAspProGlyGln 388
Qy 1710 AAAGCCTCAGTGGCTTTGGCCCTTATAGGACAGTCTGGATATTAGAAAGAGCTT 1769
Db 389 GluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGlyArgLysLeu 408
```

```
Qy 1770 CACAGACTCGAAGGTTTACAGGAGGCTGAGTTAGCTAGCTAGGAGGAGGAGAAA 1829
Db 409 GluArgLeuGluAspLysSerLysThrLeuGlyAspLeuValArgGluAlaGluArg 428
Qy 1830 GTATATTACAAAAGGAGACAGAGAAGAAAGGAAACAAAGGAAAGAGAGAGAGAG 1889
Db 429 IlePheAsnLysGlyGluThrProGluGluArgGluValArgGluThrGlu 448
Qy 1890 GAAAGGAGGAGAAAGA-----CCTAATAAACCG 1916
Db 449 GluLysGluGluArgArgArgAlaGluGluGlnLysGluLysGluArgAspArgArg 468
Qy 1917 CAAGAGAGAAATTTGACTAGATCTTGGCTGCGAGTGGTGAAGGGAAGCAATACCGAA 1976
Db 469 ArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGln----- 484
Qy 1977 AGAGAGAGAGATTTTAGGAAAATTAGTTCAGGCGCTAGACAGTCAAGGAACTGGGCAAT 2036
Db 485 -----ArgGlnAspArgGlnGlyGluArg-----Arg 494
Qy 2037 AGGACCCCACTCAAGAGGACCAATGTGCATATTGTAAAGAAAGAGACACTGGGCAAG 2096
Db 495 ArgProGlnLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaLys 514
Qy 2097 AACTGCCCCCAAGAG-----GGAACAAGGACCAAG-----ATCCTAGCT 2138
Db 515 AspCysProLysLysProArgGlyProArgGlyProArgProGlnThrSerLeuLeuThr 534
Qy 2139 CTAGAAGAA 2147
Db 535 LeuAspAsp 537

RESULT 4
US-10-510-947-6
; Sequence 6, Application US/10510947
; Publication No. US20050255123A1
; GENERAL INFORMATION:
; APPLICANT: The Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Medina, Maria Fe C.
; APPLICANT: Kobinger, Gary
; TITLE OF INVENTION: Chimeric Ebola Virus Envelopes and Uses Therefor
; FILE REFERENCE: UPN-02811PCT
; CURRENT APPLICATION NUMBER: US/10/510,947
; CURRENT FILING DATE: 2004-10-28
; PRIOR APPLICATION NUMBER: US 60/376,480
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/385,704
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/427,752
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Murine leukemia virus
US-10-510-947-6

Alignment Scores:
Pred. No.: 4.6e-81 Length: 654
Score: 1267.50 Matches: 277
Percent Similarity: 57.0% Conservative: 91
Best Local Similarity: 42.9% Mismatches: 189
Query Match: 8.7% Indels: 89
DB: 6 Gaps: 16

US-10-723-552-3 (1-8132) x US-10-510-947-6 (1-654)
Qy 5749 ATCGCATAGGACAGACGCTGAACTCCCATAAACCCCTTATCTCTACCTGGTTAACTACT 5808
Db 27 LeuArgValGlyMetAlaGluSerProHisGlnValPheAsnValThrTrpArgValThr 46
```



```

; SEQ ID NO 12
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-042-988-12

Alignment Scores:
Pred. No.: 1.49e-22 Length: 912
Score: 431.00 Matches: 209
Percent Similarity: 35.9% Conservative: 145
Best Local Similarity: 21.2% Mismatches: 345
Query Match: 2.9% Gaps: 286
DB: 7

US-10-723-552-3 (1-8132) x US-11-042-988-12 (1-912)
QY 2391 CTGGTCATACCTGAGTCCCGCAGCACCCCTCTTAGGTAGAGACTATTGACCAAGATGGGA 2450
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 40 ValLeuValGlyProThrProValAsnIleIleGlyArgAsnLeuLeuThrGlnIleGly 59
QY 2451 GCACAAATTTCTTTTGAACAAGGAAACCCAGAAAGTGTCTGCAATAACAACACTATCACT 2510
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 CysThrLeuAsnPhe-----ProIle----- 66
QY 2511 GTGTGTGACCTCCAAATTAGATGACGAATATCGACTATCTCTCCCTAGTAAGCCTGAT 2570
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 -----SerProIle----- 69
QY 2571 CAATAATATACATTCCTGGTGGAAACAGTTTCCCCCAAGCCTGGGCGAGAAACGGCAGGGATG 2630
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 ----- 69
QY 2631 GGTTTGCCAAAGCAAGTTTCCCCCAAGTTATTTCAACTGAAGGCCAGTGCACACACAGTG 2690
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 -----GluThrValPro-----ValLysLeuLysProGlyMetAspGlyPro 83
QY 2691 TCAGTCACAGCAGTACCCCTTGAGTAAAGAAGCTCAAGAAGAAATTCGCCGCGCATGTCCAA 2750
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 LysValLysGlnTrpProLeuThrGluGluLysIleLysAlaLeuValGluIleCysThr 103
QY 2751 AGATTATCCACAGGCGATCCTAGTTCCTGTC-----CAATCTCCTCGAATACTCTCC 2804
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 104 GluMetGluLysGluGlyLysIleSerLysIleGlyProGluAsnProTyrAsnThrPro 123
QY 2805 CTGCTACCGGTTAGAAAGCCTGGCGACTAATGACTATCGACAGTACAGGACTCTGAGAGAG 2864
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 ValPheAlaIleLysLysLysAspSerThrLysTrpArgLysLeuValAspPheArgGlu 143
QY 2865 GTCAATAAAGCGGTGCGAGGAT-----ATACACCCCAACAGTCCCGAACCCCTTATAAC 2915
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 LeuAsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAlaGly 163
QY 2916 CTCCTTGCTGCTCTCCACCCCAAGGAGCTGGTATACAGTATTGGACTTTAAAGGATGCC 2975
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 164 Leu-----LysLysLysLysSerValThrValLeuAspValGlyAspAla 178
QY 2976 TTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGTAATGGAGAGAT 3035
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 TyrSerValProProLeuAspGluAspPheArgLysTyrThrAlaPheThrIleProSer 198
QY 3036 CCAGGTACGGGAAGAACCGGG---CAGCTCACCTGGACCCGACCTGCCCGCAAGGGTTCAAG 3092
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 IleAsnAsnGluThrProGlyLysIleArgTyrGlnTyrAsnValLeuProGlnGlyTrpLys 218
QY 3093 AACTCCCGGACCAATCTTTTGACGAAGCCCTACACAGAGACCTTGGCCCACTTCAGGATCCAA 3152
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 219 GlySerProAlaIlePheGlnSerSerMetThrLysIleLeuGluProPheArgLysGln 238
QY 3153 CACCCTCAGGTGACCCCTCTCCAGTACGTGGATGACTGCTTCTGCGGGGAGGCCACCCAAA 3212
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 AsnProAspIleValIleTyrGlnTyrMetAspAspLeuTyrVal---GlySerAspLeu 257
QY 3213 CAGGACTGCTTAGAAGGCAAGG---GCATCTACTGCTGGAATTTGCTCTGACTAGGC 3266

```

Db	258	GiulleGlnHisArgThrLysIleGluGluLeuArgGlnHisLeuLeuArgTrpGly	277
Qy	3267	TACAGACCTCTGCTAAGAAAGGCCAGATTTCAGGAGGAGGTAAACATATTG	3320
Db	278	LeuThrThrProAspLysLysHisGln-----LysGluProProPheLeuTrpMet	294
Qy	3321	GGGTACAGTTTGGGGAGCGGACGCGATGCTGACGGAGCAGGAAGAAACTGTAGTC	3380
Db	295	GlyThrGluLeuHisPro---AspLysTrp-----ThrValGln	306
Qy	3381	CAGATACCGGCCCCACCC-----ACAGCCAAACAATGAGAGAGTTTTCGGG	3428
Db	307	ProfileValLeuProGluLysAspSerTrpThrValAsnAspIleGlnLysLeuValGly	326
Qy	3429	ACAGCTGGATTTCAGACTGTGGATCCCGGGGTTT-----	3464
Db	327	LysLeuAsnTrpAlaSerGlnIleTyProGlyIleLysValArgGlnLeuCysLysLeu	346
Qy	3465	-----OCGACCTTAGCAGCCCCACTCTACCCGCTAACCAAA-----	3500
Db	347	LeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGluGluAlaGluLeuGlu	366
Qy	3501	-----GAAAAAGGGGAATTCTCTGGGCTCCTGAGCACCAGAGGCAATTTGATCTATC	3554
Db	367	LeuAlaGluAsnArgGluIleLeuLysGluProValHisGlyValTyTyAspProSer	386
Qy	3555	AAAAAGCCCTGTCTAGCGCACCTGTCTCTGGCCCTCCCTGACGTAACTAAACCTTTACC	3614
Db	387	LysAspLeuIleAlaGluIleGlnLysGlnGlyGlnTrpThr-----TyrGln	404
Qy	3615	CTTTATGTGGATGAGCGTAAGGAGTAGCCCGGGAGTTTTAAACCCAAACCTTAGGACCA	3674
Db	405	IleTyrglnGluProPheLysAsnLeuLysThrGlyLysTyAlaArgMetArgGlyAla	424
Qy	3675	TGGAGAAGACTGTGCGCTACCTGTCAAAGAAGCTCGATCCTGTAGCAGCGTGTGTGGCCC	3734
Db	425	HisThrAsnAspValLysGlnLeuThrGluAlaValGlnLysIleThrThr-----	441
Qy	3735	ATATGCTGAAGGCTATCGCAGCTGTGGCCATCTGCTCAAGCAGCGCTGACAAATTGACT	3794
Db	442	-----GluSerIleValIleTrpGlyLys-----	449
Qy	3795	TTGGGACAGAAATAAATGTAATAGCCCCCATGCAATGGAGAACATCGTTCGGCAGCCC	3854
Db	450	-----ThrProLysPheLysLeuProIleGlnLysGluThr	461
Qy	3855	CCAGACGATGGATGACCAACGCCCGCATGACCCACTATCAAGCCTGCTTCTCACAGAG	3914
Db	462	TrpGluThrTrpTrpThrGlu-----TyrTrpGln-----	471
Qy	3915	AGGTCACGTTGCTGCCACCAGCGGCTCTCAACCCCTGCCACTCTTCCTGCCCTGAA-----	3968
Db	472	-----AlaThrTrpIleProGluTrpGlu	479
Qy	3969	---GAGACTGATGAACAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGG	4025
Db	480	PheValAsnThrProProLeuValLysLeuTrpTyrglnLeuGluLysGlu-----	496
Qy	4026	GTCGCGAAGGACCTTACAGACATACCGCTGACTGGGAAGTGCTTAACCTGGTTCACGTGAC	4085
Db	497	-----ProIleValGly---AlaGluThrPheTyValAsp	507
Qy	4086	GGAGCAGC-----TATGTGGTGGAAAGGTAAACAGG	4115
Db	508	GlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGlyTyValThrAsnArgGlyArg	527
Qy	4116	ATGCTGGGGCGGGTGGTGGAGCGGACCGCACGATCTGGGCCAGCAGCTCCCGCGAA	4175
Db	528	GlnLysValValThrLeuThrAsp-----	535
Qy	4176	GGAACTTCAGCAAAAAGGTGAGTTCATGGCCCTCACGCAAGCTTTGGCGGTGCGCGAA	4235

```
Db 536 ---ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyrLeuAlaLeuGln---AspSer 553
QY 4236 GGGAAATCCATAAATTTATACGACAGAGATGATGCTTTGCGACTGCACACGTACAT 4295
Db 554 GlyLeuGluValAlaIleValThrAspSerGlnTyrAlaLeuGly-----568
QY 4296 GGGCCCATCTATAACAAGGGGTTCTTACCTCAGCAGGGAGGGAATAAAGAACAAA 4355
Db 569 -----IleIleGlnAlaGlnProAspGlnSerGluSer 579
QY 4356 GAGGAATCTTAAGCCTATTAGACCGGTACATTTACCAAAAGGCTAGCTATTATACAC 4415
Db 580 GluLeuValAsnGlnIleIleGluGlnLeuIleLysGlnLysValTyrLeuAlaTyr 599
QY 4416 TGTCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATGGCTGACCGG 4475
Db 600 ValProAlaHisLysGlyIle-----GlyGlyAsnGluGlnValAspLys 614
QY 4476 GTTCCAGCAGGACGCCAGGCTGTAACTTCTGCTCTATATAGAAATGCCCAAGCC 4535
Db 615 LeuValSerAlaGlyIleArgLysValLeuPheLeu-----626
QY 4536 CCAGAACCCAGACGACAGTACACCTTAGAAGACTGGCAAGAGATAAAAAGATAGACCAG 4595
Db 626 -----626
QY 4596 TTCTCTGAGACTCCGGAAGGACCTGCTATATACCTCAGATGGGAAGAAATCTCTGCCCCAC 4655
Db 626 -----626
QY 4656 AAGAAGGTTAGATATGTCACACAGATACATGCTTAACCCACCTAGGAATCTAAACAC 4715
Db 627 ---AspGlyIleAspLysAlaGlnAspGluHisGlnLysTyrHisSerAsnTyrAlaGla 645
QY 4716 CTCGACGAGTGTGCAGAACATCCCTTATCATGTTCTGAGGCTACGAGAGTGGCTGAC 4775
Db 646 Met-----AlaSerAspPheAsnLeu-----ProProValValAlaLys 658
QY 4776 TCGGTGTCAACATTTGTGTGCCCTGCCAGCTGGTTAATGCTAATCTTCCAGAATGCCCT 4835
Db 659 GluIleValAlaSerCysAspLysCysGlnLeu-----669
QY 4836 CAGGAAGAGACTAAGGGGAAGC-----CACCCAGCGCTCACTGGGAAGTGGAC 4886
Db 670 LysGlyGluAlaMetHisGlyGlnValAspCysSerProGlyIle---TyrGlnLeuAsp 688
QY 4887 TTCACCTAGGTAAAGCCGGCTAATACGGNAACAATACCTATTGTTTGTGTAGACACC 4946
Db 689 CysThrHisLeuGlu-----GlyLysValIleLeuValAlaValHisVal 703
QY 4947 TTTTCAGGATGGGTAGAGGCTTATCTTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 5006
Db 704 AlaSerGlyTyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyr 723
QY 5007 AAAATACGTGAAGAAATTTTCCAAGATTTGGAAATACAGTAATAGGTCAGACAAT 5066
Db 724 PheLeuLeu---LysLeuAlaGlyArgTyrProVal---LysThrIleHisThrAspAsn 741
QY 5067 GGTCCAGCTTTTGTCCCGAGTAACTCAGGAGTGGCCAGATATATGGGATGATGG 5126
Db 742 GlySerAsnPheThrGlyAlaThrValArgAlaAlaCysTyrTyrPheAlaGlyIleLysGln 761
QY 5127 AAACCTGATGTGCATACAGACCCCAAGCTCAGGACAGGTAGAGAGGATGAATAGAAC 5186
Db 762 GluPheGlyIleProTyrAsnProGlnSerGlnGlyValValGluSerMetAsnLysGlu 781
QY 5187 ATTAAGAGACCCCTT 5201
Db 782 LeuLysLysIleIle 786
```

RESULT 6

US-10-821-234-1395

; Sequence 1395, Application US/10821234

```
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1395
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1395

Alignment Scores:
Pred. No.: 2,12e-21 Length: 540
Score: 413.50 Matches: 139
Percent Similarity: 38.5% Conservative: 75
Best Local Similarity: 25.0% Mismatches: 201
Query Match: 2.8% Indels: 141
DB: 6 Gaps: 21

US-10-723-552-3 (1-8132) x US-10-821-234-1395 (1-540)
QY 5962 GGACCAACCAATAATGGAACAAATTCGGAAATCCGAGAGATTTCTTTTGAACAATGG 6021
Db 60 GlyThrProThrPheThrAlaHisThrHisMetProArgAsnCysTyrHisSerAlaThr 79
QY 6022 AACTGTGTAACTCTTAATGATGATATGGAATGGAATGCCAACCTCTCAGCAGGATAGGTA 6081
Db 80 LeuCysMetHisAlaAsnThrHisTyrTrp-----89
QY 6082 AGTTTTTCTTATGTCACACCTATACCGCTCTGGACAATTTAATTACCTGACCTGGATT 6141
Db 89 -----89
QY 6142 AGAACTTGA-----AGCCCCAAGTGTCTCTCTTCAGACCTAGATTACCTTAAACAA--- 6189
Db 90 ---ThrGlyLysMetIleAsnProSerCys---ProGlyGlyLeuGlyValThrValCys 107
QY 6190 ---ATAAGTTTCACTGAGAAAGAAACAAAGAAATATCTCTAAATGGGTAAATGGTATG 6246
Db 108 TrpThrTyrPheThrGlnThrGlyMetSerAsp-----118
QY 6247 TCTTGGGAATGGTATATATTATGGAGGCTCGGGTAAACACAGGCTCCATCTTAACTATT 6306
Db 119 -----GlyGlyGlyValGlnAspGlnAlaArgGluLysHisVal 131
QY 6307 CGCCTCAAAATAAACCAGCTGGAGCCTCCCAATGGCTATAGGACCAATAACGGTCTTGACG 6366
Db 132 LysGluValIleSerGlnLeu-----138
QY 6367 GGTCAAGACCCCAACCAACCAAGCAGCAGCATCTCTTAACATACTTCTGTGATCAGAC 6426
Db 139 -----ThrArgValHisGlyThrSerSerProTyr---LysGlyLeuAsp 152
QY 6427 CCCACTGAGTCTAACAGCAGCACTAAATGGGGCAAACTTTTATAGCTCATCCAGGGA 6486
Db 153 LeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThr 172
QY 6487 GCTTTTCAAGCTCTTAACTCCACGACTCCAGAGGCTACCTCTCTTGTGGTATGCTTA 6546
Db 173 ThrLeuThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeu 192
QY 6547 GCTTCGGGC-----CCACTTACTATGAGGAATGGCT 6579
Db 193 ProLeuAsnPheArgProTyrValSerIleProValProGluGlnTrpAsnAsnPheSer 212
```

```
QY 6580 AGAAGA-----GGGAAATTCATGTGACA 6603
Db 213 ThrGluIleAsnThrThrSerValLeuValGlyProLeuValSerAsnLeuGluIleThr 232
QY 6604 AAAGAACATAGACCAATGCATCGGGATCCCAAAATAGCTTACCTTACTCTAGAGTT 6663
Db 233 HisThrSerAsnLeuThrCysVal-----LysPheSerAsnThrThrTyr 247
QY 6664 TCTGGAAAGCCACTGCATAGGAAGGTTCCCCCATCCACCAACACCTTTGTAAACCAC 6723
Db 248 ThrThrAsnSerGlnCysIleArgTrpValThrProThrGlnIleValCysLeuPro 267
QY 6724 ACTGAAGCCTTAATCAAACTCTGAGAGTCAATATCTGGTACCTGCTTATGACAGGTGG 6783
Db 268 SerGlyIlePhe----- 271
QY 6784 TGGGCATGTAATACTGATTAACCCCTTGTTTCCACCTTGGTTTTTAACCAAACTAAA 6843
Db 272 PheValCysGlyThrSerAlaTyrArgCysLeuAsn-----GlySerSerGlu 287
QY 6844 GATTTTGCATTATGTCCTCAAAATGTTCCCGAGTGATTACTATCCCGAAAAAGCAATC 6903
Db 288 SerMetCysPheLeuSerPheLeuValProProMetThrIleTyrThrGluGln----- 305
QY 6904 CTTGATGAATATGACTAC-----AGAAATCATCGACAAAGAGAGAACCCATATCTCTG 6957
Db 306 ---AspLeuTyrSerTyrValIleSerLysPheProArgAsnLysArgValProIle----- 322
QY 6958 ACACCTCTGCTGATGCTCGGAGTGGAGTGGCAGCGGTGTAGGAAACAGGAACAGCTGCC 7017
Db 323 ---LeuProPheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGly 341
QY 7018 CTGCTACAGGGACACAGAGCTAGAAACAGAGCTTAGTAACCTACATCGAAATGTAAACA 7077
Db 342 IleThrThrSerThrGln-----PheTyrTyrLysLeuSerGlnGluLeuAsn 357
QY 7078 GAAGATCTCCACGCCCTAGAAAATCTGTCAGTAACCTGGAGGAACTCCCTAACTCTTA 7137
Db 358 GlyAspMetGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeuAsnSerLeu 377
QY 7138 TCTGAAGTAGCTCTACAGAAATAGAGAGGGTTAGATTATTATTCTTAAAGAGGAGGA 7197
Db 378 AlaAlaValValLeuGlnAsnArgAlaLeuAspLeuLeuThrAlaGluArgGlyGly 397
QY 7198 TTATGTACGCTTGAAGGAGGAATGCTGTTTTTATGTGATCATTCAGGGGCCATCAGA 7257
Db 398 ThrCysLeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThr 417
QY 7258 GACTCCATGAACAGCTTAGAGAAAGGTTGGAGAGCGTCGAGAGGGAAGAAACTACT 7317
Db 418 GluIleValLysGluIleArgAspArgIleGlnArgAlaGluGluLeuArgAsnThr 437
QY 7318 CAAGGGTGGTTGAGGATGGTTCAACAGGTCTCTTTGGTTGGCTACCTACTTTCTGCT 7377
Db 438 GlyProTrp-----GlyLeuLeu-----SerGlnTrpMetProTrpIleLeuProPhe 453
QY 7378 TTAACAGGACCTTAATAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7437
Db 454 Leu---GlyProLeuAlaIleIleLeuLeuLeuPheGlyProCysIlePheAsn 472
QY 7438 AGGTAAATTCCTTATTAGAGAACGAATAAGTGCAGTCCAGATC---ATGGTACTTAGA 7494
Db 473 LeuLeuValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMetGluProLys 492
QY 7495 CAACAG-----TACCAAGCCCTCTAGCAGGGAAGCT 7527
Db 493 MetGlnSerLysThrLysIleTyrArgArgProLeuAspArgProAla 508
```

RESULT 7

US-10-507-928-10

; Sequence 10, Application US/10507928

; Publication No. US20050266024A1

```
; GENERAL INFORMATION:
; APPLICANT: POWDERMED LIMITED AND GLAXO GROUP LIMITED
; TITLE OF INVENTION: ADJUVANT
; FILE REFERENCE: N.88232B GCW
; CURRENT APPLICATION NUMBER: US/10/507,928
; CURRENT FILING DATE: 2004-09-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of RT insert of p7077-RT3
US-10-507-928-10
```

Alignment Scores:

Pred. No.:	1-56e-13	Length:	562
Score:	301.00	Matches:	145
Percent Similarity:	37.2%	Conservative:	104
Best Local Similarity:	21.6%	Mismatches:	267
Query Match:	2.1%	Indels:	154
DB:	6	Gaps:	26

US-10-723-552-3 (1-8132) x US-10-507-928-10 (1-562)

```
QY 2628 ATGGGTTTGGCAAGCAAGTTCCCCACAAAGTTATTCAACTGAAGGCCAGTGCACACCA 2687
Db 1 MetGlyProIleSerProIleGluThrValSerValLysLeuLysProGlyMetAspGly 20
QY 2688 GTGTCTAGTCAGACAGTACCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCCGCGATGC 2747
Db 21 ProLysValLysGlnTrpProLeuThrGluGluLysIleLysAlaLeuValGluIleCys 40
QY 2748 CAAAGATTATATCCAACAGGCGCATCTAGTTCTCTGTC-----CAATCTCCCTGGAATFAC 2801
Db 41 ThrGluMetGluLysGluGlyLysIleSerLysIleGlyProGluAsnProTyrAsnThr 60
QY 2802 CCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCACTAGCAGGACTTGAGA 2861
Db 61 ProValPheAlaIleLysLysLysAspSerThrLysTrpArgLysLeuValAspPheArg 80
QY 2862 GAGGTCAATAAAGCGGTGCAGGAT-----ATACACCAACAGTCCCGAACCCCTTAT 2912
Db 81 GluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAla 100
QY 2913 AACCTTGTGTGCTCTCCACCCCAACCGAGCTGTGTATACAGTATTGGACTTAAAGGAT 2972
Db 101 GlyLeu-----LysLysLysSerValThrValLeuAspValGlyAsp 115
QY 2973 GCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCTGATGGAGA 3032
Db 116 AlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIlePro 135
QY 3033 GATCCAGGTACGGGAACCCGGG---CAGCTCACCTGACCCGACTGCCCCCAAGGGTTC 3089
Db 136 SerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGlyTrp 155
QY 3090 AAGAACTCCCGACCACTTTTGAAGCCCTACACAGACCTGCGCAACTTCAGGATC 3149
Db 156 LysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeuGluProPheArgLys 175
QY 3150 CAACACCTCAGGTGACCTCTCCAGTACGTGATGACCTGTTCTGGCGGAGGCCACC 3209
Db 176 GlnAsnProAspIleValIleTyrGlnTyrMetAspLeuTyrVal---GlySerAsp 194
QY 3210 AAACAGACTGCTTAGAAGGCACGAAG-----GCACACTGCTGGAATGTCTGACCTA 3263
Db 195 LeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGlnHisLeuLeuArgTrp 214
QY 3264 GGCTACAGAGCCTCTCTCTAAGAGGCCCAAGATTTGCGAGGAGAGAGGTAAACATCTTG--- 3320
Db 215 GlyLeuThrThrProAspLysLysHisGln-----LysGluProProPheLeuTrp 231
```



```
QY 3033 GATCCAGGTACGGGAACCGGG---CAGCTCACCTGGACCGGACCTGCCCGGAGGTTTC 3089
Db 136 SerIleAsnAenGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGlyTrp 155
QY 3090 AAGAACTCCCGGACCATCTTTGACGAAGCCCTACACAGACACCTGGCCAACTTCAGGATC 3149
Db 156 LysGlySerProAlaIlePheGlnSerMetThrLysIleLeuGluProPheArgLys 175
QY 3150 CAACACCTCAGGTGACCCCTCCAGTACGTGATGACCTGCTTTCGGCGAGGCCACC 3209
Db 176 GlnAsnProAspIleValIleTyrGlnTyrMetAspLeuTyrVal---GlySerAsp 194
QY 3210 AAACAGCACTGCTTAGAAGCAGCAG---GCACACTGCTGGNAATTGCTCGACCTA 3263
Db 195 LeuGluIleGlyGlnHisArgThrLysIleGluLeuArgGlnHisLeuLeuArgTrp 214
QY 3264 GGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTCAGAGAGAGAGTAACATACTTG--- 3320
Db 215 GlyLeuThrThrProAspLysHisGln-----LysGluProProPheLeuTrp 231
QY 3321 ---GGGTACATTTGCGGACGGGACGATGCTGACGAGCGGACGGAAGAAACTGTA 3377
Db 232 MetGlyTyrGluLeuHisPro---AspLysTrp-----ThrVal 243
QY 3378 GTCCAGATACGGCCCCCAACC-----ACAGCCAAACAAATGACAGAGTTTGTG 3425
Db 244 GlnProIleValLeuProGluLysAspSerTrpThrValAsnAspIleGlnLysLeuVal 263
QY 3426 GGGACAGCTGATTTGCAGACTGTGGATCCCGGGGTTT----- 3464
Db 264 GlyLysLeuAsnTrpAlaSerGlnIleTyrProGlyIleLysValArgGlnLysLys 283
QY 3465 -----GCGACCTTAGCAGCCCCCACTTACCCGCTAACCAAA----- 3500
Db 284 LeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGluGluAlaGluLeu 303
QY 3501 -----CAAAAGGGGATTTCTCTGGCTCTGAGCAGCAGGAGGATTTGATGCT 3551
Db 304 GluLeuAlaGluAsnArgGluIleLeuLysGluProValHisGlyValTyrTrpAspPro 323
QY 3552 ATCAAAAGGCCCTGTGAGCGCACCTGTCTGCGCCCTCCCTGACGTAACTAAACCCCTTT 3611
Db 324 SerLysAspLeuIleAlaGluIleGlnLysGlnGlyGlnTrpThr-----Tyr 341
QY 3612 ACCCTTTATGTGATGAGCTAGAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGA 3671
Db 342 GlnIleTyrGlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArgMetArgGly 361
QY 3672 CCATGGAGAAGACTGTGCGCTACCTGTCAAGAGCTCGATCCTGTAGCCAGTGGTGG 3731
Db 362 AlaHisThrAsnAspValLysGlnLeuThrGluAlaValGlnLysIleThrThr----- 379
QY 3732 CCCATATGCTGAAGGCTATCGCAGCTGTGGCCACTGTGTCAGGACGCTGACAAATTG 3791
Db 380 -----GluSerIleValIleTrpGlyLys----- 387
QY 3792 ACTTTGGGACAGATATACTGTAATAGCCCCCATGCAATGCGAGACATCGTTCCGCAG 3851
Db 388 -----ThrProLysPheLysLeuProIleGlnLysGlu 398
QY 3852 CCCCACACCGATGATGACCAACGCCCGGATGACCCACTATCAAGCCCTGCTCTCACA 3911
Db 399 ThrTrpGluThrTrpTrpTrpGlu-----TyrTrpGln----- 409
QY 3912 GAGAGGGTCAAGTTGCTGCCACAGCGCGCTCTCAACCCCTGCCACTCTTCTGCTGAA--- 3968
Db 410 -----AlaThrTrpIleProGluTrp 416
QY 3969 -----GAGACTGATGAACGAGTACTCATGATTGCGCATCAACTATTGATGAGGAGACT 4022
Db 417 GluPheValAsnThrProProLeuValLysLeuTrpTyrGlnLeuGluLysGlu----- 434
```

```
QY 4023 GGGGTCCGCAAGGACCTTTACAGACATACCGCTGACTGGAGAGTGTCTAACTGGTTCACT 4082
Db 435 -----ProIleValGly---AlaGluThrPheTyrVal 444
QY 4083 GACGGAAGCAGC-----TATGTGGTGAAGGTAAG 4112
Db 445 AspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGlyTyrValThrAsnArgGly 464
QY 4113 AGATGGCTGGGGCGGGTGGTGGAGCGGACCGCACGATCTGGGCGCAGCAGCTGCCG 4172
Db 465 ArgGlnLysValValThrLeuThrAsp----- 473
QY 4173 GAAGGAATCTCAGCACAAAAGGCTGAGCTCATGGCCCTCAGCAAGCTTTGGCGTGGCC 4232
Db 474 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyrLeuAlaLeuGln--- 490
QY 4233 GAAGGAATCTCATAAATTTATACGACAGCAGGATGCTTTCGACATGCACACGTA 4292
Db 491 SerGlyLeuGluValAsnIleValThrAspSerGlnTyrAlaLeuGly----- 506
QY 4293 CATGGGGCCATCTATAAACAAGGGGTGCTTACCTCAGCAGGAGGGAATAAAGAAC 4352
Db 507 -----IleIleGlnAlaGlnProAspGlnSerGlu 516
QY 4353 AAAGAGGAATCTTAAGCCTATTAGAAGCGGTACATTTCACAAAAGGCTAGCTATTATA 4412
Db 517 SerGluLeuValAsnGlnIleIleGluGlnLeuIleLysGluLysValTyrLeuAla 536
QY 4413 CACTGTCTGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAACCAAGATGGCTGAC 4472
Db 537 TrpValProAlaHisLysGlyIle-----GlyGlyAsnGluGlnValAsp 551
QY 4473 CGGGTTCAGCAGCAGGCGCCAGGGTGT 4502
Db 552 LysLeuValSerAlaGlyIleArgLysVal 561

RESULT 9
US-11-029-465-10
; Sequence 10, Application US/11029465
; Publication No. US20050256070A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Ralph P.
; APPLICANT: Thomsen, Lindy
; APPLICANT: Van-Wely, Catherine
; APPLICANT: Ertl, Peter
; TITLE OF INVENTION: Adjuvant
; FILE REFERENCE: 033267-015
; CURRENT APPLICATION NUMBER: US/11/029,465
; CURRENT FILING DATE: 2005-01-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of RT insert of p7077-RT3
US-11-029-465-10

Alignment Scores:
Pred. No.: 1,56e-13 Length: 562
Score: 301.00 Matches: 145
Percent Similarity: 37.2% Conservative: 104
Best Local Similarity: 21.6% Mismatches: 267
Query Match: 2.1% Indels: 154
DB: 7 Gaps: 26

US-10-723-552-3 (1-8132) x US-11-029-465-10 (1-562)
QY 2628 ATGGGTTTGGCAAGCAGTTCCCCACAAAGTTATTTCACACTGAAGCCAGTGCACACCA 2687
Db 1 MetGlyProIleSerProIleGluThrValSerValLysLeuLysProGlyMetAspGly 20
```



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of the coding insert in
; OTHER INFORMATION: p731-RT3
US-11-029-465-12
```

```
Alignment Scores:
Pred. No.: 1.37e-12 Length: 561
Score: 287.50 Matches: 139
Percent Similarity: 36.1% Conservative: 105
Best Local Similarity: 20.8% Mismatches: 265
Query Match: 2.0% Indels: 167
DB: 7 Gaps: 24
```

US-10-723-552-3 (1-8132) x US-11-029-465-12 (1-561)

```
QY 2628 ATGGGTTGGCAAGCAAGTTCCCCCAACAGTTATTCACTGAAGGCCAGTGCACACCA 2687
D 1 MetGlyProIleSerProIleGluThrValSerValLysLeuLysProGlyMetAspGly 20
QY 2688 GTGTCAGTCAGACAGTACCCCTTCAGTAAAGAAAGCTCAAGAAAGAAATTCGGCCGCATGTC 2747
D 21 ProLysValLysGln---ProLeuThrGluGluLysLysLysLysAlaLeuValGluLeuCys 39
QY 2748 CAAGATTAAATCCCAACAGGCGATCTAGTTCTCTGTC-----CAATCTCCCTGGGAATACT 2801
D 40 ThrGluMetGluLysGluGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLys 59
QY 2802 CCCCTGCTACCGGTTAGAACCCCTGGGACTAATGACTATCGACACGACGACGACGACGACGACG 2861
D 60 ProValPheAlaIleLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 79
QY 2862 GAGGTCAATAAAGCGGTGCAGAT-----ATACACCCCAACAGTCCCGAACCCCTTAT 2912
D 80 GluLeuAsnLysArgThrGlnAspPheThrGluValGlnLeuGlyLysLysLysLysLysLys 99
QY 2913 AACCTCTGTGTGCTCTCCCAACCGGAGCTGGGTATACAGTATTGGACTTAAAGGAT 2972
D 100 GlyLeu-----LysLysLysLysSerValThrValLeuAspValGlyAsp 114
QY 2973 GCCTTCTCTGCTGAGATTAACCCCACTAGCAACCACTTTTGCCTTCGATGGAGA 3032
D 115 AlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIlePro 134
QY 3033 GATCAGGTACGGGAAGAACCGG---CAGCTCACCTGGACCCGACTGCCCAAGGGTTC 3089
D 135 SerIleAsnAsnGluThrProGlyLysLeuArgTyrGlnTyrAsnValLeuProGlnGlyTyr 154
QY 3090 AAGAACTCCCGACCATTTTGACGAAGCCCTACACAGACGACCTGGCCAACTTCAGGATC 3149
D 155 LysGlySerProAlaIlePheGlnSerSerMetThrLysLysLysLysLysLysLysLysLys 174
QY 3150 CAACACCTCAGGTGACCCCTCTCCAGTACGTGATGACCTGCTCTGGCGGGAGCCACC 3209
D 175 GlnAsnProAlaIleValIleTyrGlnTyrMetAspLysLeuTyrValGly----- 191
QY 3210 AAACAGGACTCTGTAGAAGGCACGAAGGCACTACTGCTGGAATTTGTCTGACTAGGCTAC 3269
D 192 -----SerAspLeuGluLeu 196
QY 3270 AGAGCCTCTGTAAGAAGGCCCAAGATTTCAGGAGAGAGGTAAACATCTTGGGGTACAGT 3329
D 197 GlyGlnHisThrArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 216
QY 3330 TTGGCGGAGCGGACGCGA-----TGCTG-----ACGAG 3359
D 217 ThrProAspLysLysHisGlnLysGluProProPheLeuTyrMetGlyTyrGluLeuHis 236
QY 3360 GCACGGAAGAAAACCTGTAGTCCAGATACCGGCCCAACC-----ACAGCCAAA 3407
```

```
D 237 ProAspLysTyrThrValGlnProIleValLeuProGluLysAspSerTyrThrValAsn 256
QY 3408 CAAATGAGAGAGTTCCTGGGAGCAGCTGGATTTTGCAGACTGCTGATCCCGGGGTTT--- 3464
D 257 AspIleGlnLysLeuValGlyLysLeuAsnTyrPalaSerGlnIleTyrProGlyLys 276
QY 3465 -----GCGACCTTAGCAGCCCTCTACCCGCTA 3494
D 277 ValArgGlnLeuCysLysLeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeu 296
QY 3495 ACCAAA-----GAAAAAGGGGAATTCCTCGGCTCTCTGAGCAC 3533
D 297 ThrGluGluAlaGluLeuGluAlaGluAsnArgGluLeuLysGluProValHis 316
QY 3534 CAGAGGCATTTGATGCTATCAAAAAGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCT 3593
D 317 GlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGlnGly 336
QY 3594 GACGTAACTAAACCTTTTACCTTTATGTGATGAGCGTAAGGAGTAGCCCGGGAGTT 3653
D 337 GlnTyrThr-----TyrGlnIleTyrGlnGluProPheLysAsnLeuLysThrGlyLys 354
QY 3654 TTAACCCCAACCTTAGGACCATGAGAGACCTGTGCGCTACCTGTCAAGAAAGCTCGAT 3713
D 355 TyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeuThrGluAlaValGln 374
QY 3714 CTTGTAGCCAGTGGTGGCCCATATGCTGAGGCTATGCGAGCTATGCGAGCTGGCCATACTGTC 3773
D 375 LysIleThrThr-----GlusIleValIleTyrGly 385
QY 3774 AAGGACGCTGACAAATTGACTTTTGGGACAGATATATACTGTAAATAGCCCCCATGCTG 3833
D 386 Lys-----ThrProLysPheLys 391
QY 3834 GAGAACATCTGTCGGCAGCCCCCAGACCGATGGATGACCAACGCCGATGACCCACTAT 3893
D 392 LeuProIleGlnLysGluThrTyrGluThrTyrTyrTyrGlu-----TyrTyr 407
QY 3894 CAAAGCCTGCTTCTCACAGAGAGGTCAGTTCGCTCCACCGCGGCTCTCAACCCCTGCC 3953
D 408 Gln-----Ala 409
QY 3954 ACTCTTCTGCTGAA-----GAGACTGTAGAACAGTGCATCATGTTGCCATCAA 4004
D 410 ThrTyrIleProGluTyrGluPheValAsnThrProProLeuValLysLeuTyrTyrGln 429
QY 4005 CTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAA 4064
D 430 LeuGluLysGlu-----ProIleValGly--- 437
QY 4065 GTGCTAACCTGGTTCACCTGACGGAAGCAGC----- 4094
D 438 AlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGly 457
QY 4095 TATGTGTGAAGGTAAAGAGGATGGCTGGGCGCGCTGGTGACGCGGACGCGACCGCACGATC 4154
D 458 TyrValThrAsnArgGlyArgGlnLysValValThrLeuThrAsp----- 472
QY 4155 TGGGCCAGCAGCTGCCGGAAGAACTTCAGCACAAAGGCTGAGCTCATGGCCCTCAGC 4214
D 473 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyr 484
QY 4215 CAAGCTTTGGGTCGGCGGAAGGAATCCATAACATTTATACGGACGACGAGTATGCC 4274
D 485 LeuAlaLeuGln---AspSerGlyLeuGluValAsnIleValThrAspSerGlnTyrAla 503
QY 4275 TTTGCGACTGACACGTCATACGCGCCATCTATAAACAAAGGGGGTGTCTTACCTCAGCA 4334
D 504 LeuGly-----IleIleGlnAla 509
QY 4335 GGGAGGGGAATAAAGAACAAAGAGGAAATTTCTAAGCCCTATTAGAAGCGGTACTTATCCA 4394
```

```
Db 510 GlnProAspGlnSerGluSerGluLeuValAsnGlnIleleGluGlnLeuLeuLysLys 529
QY 4395 AAAAGGCTAGCTATTATACATCTGCTCGACATCAGAAAGCTAAAGATCTCATATCCAGA 4454
Db 530 GlulysValTyrLeuAlaTyrValProAlaHisLysGlyLe-----Gly 544
QY 4455 GGAACACAGATGGCTGACCGGTTGCCAGCAGCAGCCCGGGTGT 4502
Db 545 GlyAsnGluGlnValAspLysLeuValSerAlaGlyIleArgLysVal 560

RESULT 11
US-11-014-629-1
; Sequence 1, Application US/11014629
; Publication No. US20050244376A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Frederick L.
; TITLE OF INVENTION: TARGETED GENE DELIVERY TO EXTRACELLULAR MATRIX
; FILE REFERENCE: 30863-704.302
; CURRENT APPLICATION NUMBER: US/11/014,629
; CURRENT FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: US 09/837,223
; PRIOR FILING DATE: 1997-04-10
; PRIOR APPLICATION NUMBER: US 09/904,923
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 229
; TYPE: PRT
; ORGANISM: moloney murine leukemia virus
US-11-014-629-1

Alignment Scores:
Pred. No.: 4.35e-10 Length: 229
Score: 250.00 Matches: 70
Percent Similarity: 39.1% Conservative: 23
Best Local Similarity: 29.4% Mismatches: 75
Query Match: 1.7% Indels: 70
DB: 1.7% Gaps: 12

US-10-723-552-3 (1-8132) x US-11-014-629-1 (1-229)
QY 5776 CATAAACCCCTATCTCTACCTGGTTAATTACTGACTCCGCGACAGTATTAATATCAAC 5835
Db 8 HisGlnValTyrAsnIleThrTrpGluValThr---AsnGlyAspArgGluThrValTrp 26
QY 5836 AACACTCAAGGCGGAGCTCTTTAGGAACCTGTGGCTGATCTATAC-----GTTTGC 5889
Db 27 AlaThrSerGlyAsnHisProLeuTrpTrpTrpProAspLeuThrProAspLeuCys 46
QY 5890 CTC-----AGATCAGTTATTCCTAGT 5910
Db 47 MetLeuAlaHisHisGlyProSerTyrTrpGlyLeuGluTyrGlnSerProPheSerSer 66
QY 5911 CTGACCTCACCCCA----- 5925
Db 67 ProProGlyProProCysSerGlySerProGlyCysSerArgAspCysGlu 86
QY 5925 ----- 5925
Db 87 GluProLeuThrSerLeuThrProArgCysAsnThrAlaTrpAsnArgLeuLysLeuAsp 106
QY 5926 GATATCTCCATGCT-----CACGGATTTTATGTTTGGCCAGGACCA-----CCAAAT 5973
Db 107 GlnThrThrHisLysSerAsnGluGlyPheTyrValCysProGlyProHisArgProArg 126
QY 5974 AATGGAACCATTCGCGAAATCCAGAGATTTCTTTTGTAAACAATGGAAGTGTGAACC 6033
Db 127 GluSerLysSerCysGlyGlyProAspSerPheTyrCysAlaTyrTrpGlyCysGluThr 146
QY 6034 TCTAATGATGATATTGGAAATGGCAACCTCTCAGCAGATAGGTAAGTTTCTTAT 6093
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
Db 147 ThrGlyArgAlaTyrTrpLys---ProSerSerSerTrpAsp-----PheIleThr 162
QY 6094 GTCAACACACCTATACCACTCTGACCAATTTAATTACCTGACCTGGATTAGAACTGGAGGC 6153
Db 163 ValAsnAsnAsnLeuThrSerAspGln-----AlaValGlnValCysLysAspAsn 179
QY 6154 CCAAGTGTCTCTCTCAGACCTAGATTACTTAAATAAGTTTCACTAGAGAAGGAAAA 6213
Db 180 LysTrpCysAsnPro-----LeuValIleArgPheThrAspAlaGlyArg 194
QY 6214 CAAGAAATATCTCTAAATGGTAAATGGTATGCTCTGGGAATGGTATATTATGGAGGC 6273
Db 195 Arg-----ValThrSerTrpThrThrGlyHisTyrTrpGlyLeuArgLeuTyr---Val 211
QY 6274 TCGGTAACAACACCAAGCTCCATTTCTAACTATTTCGCTCAAAATAAACACGAGCTG 6327
Db 212 SerGlyGlnAspProGlyLeuThrPheGlyIleArgLeuArgTyrGlnAsnLeu 229

RESULT 12
US-11-022-562-221
; Sequence 221, Application US/11022562
; Publication No. US20050249742A1
; GENERAL INFORMATION:
; APPLICANT: Ruprecht, Ruth M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; FILE REFERENCE: DFN-043CN
; CURRENT APPLICATION NUMBER: US/11/022,562
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: PCT/US03/20322
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392718
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 31, 97, 140, 141, 144, 178
; OTHER INFORMATION: Xaa = Any Amino Acid
US-11-022-562-221

Alignment Scores:
Pred. No.: 7.91e-08 Length: 210
Score: 217.50 Matches: 65
Percent Similarity: 47.9% Conservative: 37
Best Local Similarity: 30.5% Mismatches: 82
Query Match: 1.5% Indels: 29
DB: 1.5% Gaps: 9

US-10-723-552-3 (1-8132) x US-11-022-562-221 (1-210)
QY 2784 CAATCTCCCTGGAATCTCCCTGCTACCGTTAGAAAGCTGGACTAATGACTATCGA 2843
Db 16 GluAsnProTyrAsnThrProValPheAlaIleLysLysLysAsp***ThrLysTrpArg 35
QY 2844 CCAGTACAGAGCTTGAGAGAGGTCAATAAACGGGTGCAGGAT-----ATACACCCA 2894
Db 36 LysLeuValAspPheArgGluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeu 55
QY 2895 ACAGTCCCGAACCTTATAACCTTGTGTCTCTCCACCCCAACGAGAGCTGGTATACA 2954
Db 56 GlyIleProHisProAlaGlyLeu-----LysLysLysLysSerValThr 70
QY 2955 GTATTGGACTTAAAGATGCTCTTCTCTGCTGAGATTACACCCCACTAGCAACCACTT 3014
Db 71 ValLeuAspValGlyAspAlaTyrPheSerValProLeuAspGlnSerPheArgLysTyr 90
QY 3015 TTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGG---CAGCTCACCTGGACC 3071
```

```
Db      91 ThrAlaPheThrIlePro***ThrAsnAsnGluThrProGlyIleArgTyrGlnTyrAsn 110      :
Qy      3072 CGACTGCCCAAGGTTTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAC 3131      :
Db      111 ValLeuProGlnGlyTyrPheGlySerProAlaIlePheGlnSerSerMetThrLysIle 130      :
Qy      3132 CTGCCCAACTTCAGATCCACACCCCTCAGGTGACCTCTCCAGTACCTGGATGACCTG 3191      :
Db      131 LeuGluProPheArgIleLysAsnPro*****ValIle**GlnTyrMetAspAspLeu 150      :
Qy      3192 -----CTTCTGCGCGGAGCCACCAAAAGAGGACTGTTAGAAAGGCACCAAG 3236      :
Db      151 TyrValGlySerAspLeuGluIleGlyGlnHisArgAlaLysIleGluGluLeuArgLys 170      :
Qy      3237 GCATCTACTGTGGAAATGTCTGACTAGGTACAGACGCTCTGCTAAGAGAGGCCAGATT 3296      :
Db      171 HisLeuLeu-----SerTrpGlyPhe***ThrProAspLysLysHisGln--- 185      :
Qy      3297 TGCAGGAGAGAGGTAACTACTTGC-----GGGTACAGTTTGGCGGACGGCGGCGATGG 3350      :
Db      186 -----LysGluProProPheLeuTrpMetGlyTyrGluLeuHisPro---AspLysTrp 202      :
Qy      3351 CTGACGAGGAGCAGCAAGAAACTGTAGTCCAGATACCG 3389      :
Db      203 ThrValGlnPro-----IleGlnLeuPro 210      :
RESULT 13
US-11-230-251-26
; Sequence 26, Application US/11230251
; Publication No. US20060019322A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; APPLICANT: Cafferey, Robert
; APPLICANT: Ali, Shujath
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
; TITLE OF INVENTION: Genes
; FILE REFERENCE: DEX-0239
; CURRENT APPLICATION NUMBER: US/11/230,251
; CURRENT FILING DATE: 2005-09-19
; PRIOR APPLICATION NUMBER: US/09/957,708
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,746
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-230-251-26

Alignment Scores:
Pred. No.:      3,03e-07      Length:      422
Score:          210.50      Matches:      88
Percent Similarity: 39.3%      Conservative: 52
Best Local Similarity: 24.7%      Mismatches: 109
Query Match:      1.4%      Indels:      107
DB:              7          Gaps:      18

US-10-723-552-3 (1-8132) x US-11-230-251-26 (1-422)
Qy      4188 CAAAGGCTGAGCTCATGGCCCTCAGCAAGCTTTGGGTGGCGGAGGGAATCCATA 4247
Db      14 GlnArgAlaGluLeuValAlaValIleThrValLeuGln---AspPheAsnGlnSerIle 32
Qy      4248 AACATTTATACGACAGCAGGTATGCCTTT---CGGACTGCACAGCTACATGGGCCATC 4304
Db      33 AsnIleValSerAspSerAlaTyrValValGlnAlaThrLysAspIleGluArgAlaLeu 52
Qy      4305 TATAACAAAGGGGGTGTCTTACCTCAGCAGGAGGGAATATAAGAACAAAGAG----- 4358
```

```
Db      53 IleIysTyr-----IleMetAspAspGlnLeuAsn 62
Qy      4359 GAAATTTCTAAGCCTATTAGAGCCGTACATTATTACCAAAAAGG-----CTAGCTATT 4409
Db      63 ProLeuPheAsnLeuLeuGlnGln---AsnValArgLysArgAsnPheProPheTyrIle 81
Qy      4410 ATACACTGTCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACACAGATCGCT 4469
Db      82 ThrHisIleArgAlaHisThrAsnLeuProGlyProLeuThrArgAlaAsnGluGlnAla 101
Qy      4470 GACCGGGTTTCCCAAGCAGGCGACCCAGGGTGTTAACCTTCTGCTTATAATAGAAAATGCC 4529
Db      102 AspLeuLeuValSerSerAla----- 108
Qy      4530 AAAGCCCCAGAACCCAGACAGTACACCTAGAGACTGGCAAGAGATAAAAAAGATA 4589
Db      108 ----- 108
Qy      4590 GACCAGTTTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATCCTG 4649
Db      109 -----PheMetGlu----- 111
Qy      4650 CCCCACAAAGAAGGGTTAGAAATATGTCACACAGATACATCGTCTAACCCAC----- 4700
Db      112 -----AlaGlnGluLeuHisAlaLeuThrHisValAsnAla 123
Qy      4701 CTAGGAACCTAAACACACCTGCAGCAGTGGTCAGAACATCCCTTATCATGTCTCGAGCTA 4760
Db      124 IleGlyLeuLysAsnLysPheAspIleThrTrpLysGln----- 136
Qy      4761 CCAGGAGTGGCTGACTCGGTGGTCAAAACATTGTGTGCCCTGCCAGCTGGTGAAT----- 4814
Db      137 -----ThrLysAsnIleValGlnHisCysThrGlnCysGlnIleLeuHisLeuAla 153
Qy      4815 GCTAATCTTCCAGAAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGGCGCTCAC 4874
Db      154 ThrGlnGluAlaArgValAsnPro-----ArgGlyLeuCysProAsnValLeu 169
Qy      4875 TGGAGAGTGCATCTTACTGAGGTAAAGCCGCTTAATACGGAACAATACTATTCTGTT 4934
Db      170 TrpGlnMetAspValMetHisVal---ProSerPheGlyLysLeuSerPheValHisVal 188
Qy      4935 TTTGTAGACACCTTTTCAGGATGGTAGAGCT---TATCCTACTAAGAAAGAGACTTCA 4991
Db      189 ThrValAspThrTyrSerHisPheIleTrpAlaThrCysGlnThrGlyGluSerThrSer 208
Qy      4992 ACCGTGTGCTTAAAAAATACTCGAAGAAATTTTCCAGATTTGGAATCTCTTAAGGTA 5051
Db      209 HisVal-----LysArgHisLeuLeuSerCysPheProValMetGlyValProGluLys 226
Qy      5052 ATAGGTCACACAATGTCAGCTTTTGTGGCCAGGTAAAGTTCAGGACTGGCCCAAGATA 5111
Db      227 ValLysThrAspAsnGlyProGlyTyr-----CysSerLysAlaValGlnLysPhe 243
Qy      5112 TTGGGGATTGATTGGAAA-----CTGCATTGTGCATACAGACCCCAAGCTCA 5159
Db      244 LeuAsn---GlnTrpLysIleThrHisThrIleGlyIleLeuTyrAsnSerGlnGlyGln 262
Qy      5160 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGACCTTACTATAA 5207
Db      263 AlaIleIleGluArgThrAsnArgThrLeuLysAlaGlnLeuValLys 278
RESULT 14
US-11-124-367A-444
; Sequence 444, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
```


Db 760 uGluArgProAlaProProArgPheHisThrSerThrTyraSpLeuProGlyProGluGl 780
 Qy 2943 AGCTGGT-----ATACAGTATGGACTTAAAGAGTGCCTTCTTCCTGCTGAGA 2990
 Db 780 yAlaGlyProPheGluAlaSerGlnProAlaAspSerAlaValProAlaThrSerGlyLy 800
 Qy 2991 TTAC-----ACCCCACTAGCAACCACTTTTCCCTCGAAT-----GGAGAGAT 3035
 Db 800 sValTyxGlyThrProGluThrGluLeuThrTyProThrAsnIleValAlaAlaProLe 820
 Qy 3036 CCAGGTACGGGAACACCGGCACCTCACCTGGA---CCCGACTGCCCAAGGTTTCAAG 3092
 Db 820 uAlaGluGluHisValSerSerAlaThrSerIleThrGluCysAspLysLeuSerSe 840
 Qy 3093 AACTCCCCGA-----CCATCTTTGACGAAGCCCTACACAGA 3128
 Db 840 rPheAlaThrSerValAlaGluAspGlnSerValAlaSerLeuThrAlaProGlnThrGl 860
 Qy 3129 GACCTGGCCAACTTCAGATCCACACCCCTC-----AGGTACCCCTC 3170
 Db 860 u-----GluThrGlySerSerLeuLeuLeuAspThrValThrSerIleProSe 877
 Qy 3171 CTCCAGTACG-----TGTGACCTGTCTTGGGGGAGGCCACC 3209
 Db 877 rSerArgThrGluAlaThrGlnGlyLeuAspTyxVal-ProSerAlaGlyThrIleSerP 897
 Qy 3210 AAACAGGACTGCTTAGAAGGCACGAAGCACTACTGCTGGAATGTCTGACCTAGGTAC 3269
 Db 897 roThrSerSerLeuGluAspLysGlyPheLysSer-----ProProCysG 913
 Qy 3270 AGACCTCTGTAGAGGCCCACTTTGCAAGGAGAGGCTTAACATACTTGGGTACAGT 3329
 Db 913 luAspPheSerValThrGly---GluSerGluLysArgGlyGluIleGlyLysGlyL 932
 Qy 3330 TTGGCGGACGGCAGCGAGCTGCTCAC----- 3355
 Db 932 euSerGlyGluArgAlaValGluGluGluGluGluThrAlaAsnValGluMetSerG 952
 Qy 3355 ----- 3355
 Db 952 luLysLeuCysSerGlnTyxGlyThrProValPheSerAlaProGlyHisAlaLeuHisP 972
 Qy 3356 -----GGAGCCAGGAGAAACTCT---AGTCCAGAT----- 3385
 Db 972 roGlyGluProAlaLeuGlyGluAlaGluGluArgCysLeuSerProAspAspSerThrV 992
 Qy 3386 -----ACCGGCCCC----- 3394
 Db 992 alLysMetAlaSerProProSerGlyProProSerAlaThrHisThrProPheHisG 1012
 Qy 3395 -----AACCACAGCCAAACAAATGAGAGAGATTTTGGGACAGCTGGATTT 3440
 Db 1012 lnSerProValGluGluLysSerGluProGlnAspPheGlnGluAlaAspSerTrpGlyA 1032
 Qy 3441 TGCAGACTGGATCCCGGGTTGC-----GACCTTAGCAGCCCACTTACCCCTA 3494
 Db 1032 spThrLysArgThrProGlyValGlyLysGluAspAlaAlaGluGluThrValLysProG 1052
 Qy 3495 ACCAAAGAAAAGGGGAATTCTCC-----TGGGCTCCTGAG 3530
 Db 1052 ly-ProGluGluGlyThrLeuGluLysGluLysValProProProArgSerProGln 1071
 Qy 3531 CACCAGAGGCAATTGATGCTATCAAAAGCCCTGTGAGCGCAGCTGCTGCGCCCTC 3590
 Db 1072 AlaGlnGluAlaProValAsnIleAspGluGlyLeuThrGlyCysThrIleGlnLeuLeu 1091
 Qy 3591 CCTGACGTAACTAACCTTTACCTTTATGTGATGAG----- 3629
 Db 1092 ProAlaGlnAspLysAlaIleValPheGluIleMetGluAlaGlyGluProThrGlyPro 1111
 Qy 3630 -----CCTAAGGAGTAGCCCGGGAGTTTAAACCCCAACCTTAGGACCATGAGA 3680
 Db 1111 ----- 3680

Db 1112 IleLeuGlyAlaGluAlaLeuProGlyGlyLeuArgThrLeuProGlnGluProGlyLys 1131
 Qy 3681 -----AGACCTGCGCTACTCTCAAGAGCTTCGATCTCTGTAGCCAGTGGT 3728
 Db 1132 ProGlnLysAspGluValLeuArgTyxProAspAspSerLeuSerPro----- 1147
 Qy 3729 TGGCCCATATGCTGAAGGCTATCGCAGCTGTGCCCATACTGGTCAAGGAGCGCTGACAAA 3788
 Db 1148 ----- 3788
 Qy 3789 TTG-----ACTTTGGGACAGATATAACTGTATATAGCC 3821
 Db 1153 LeuSerValLeuSerValProSerProAspThrAlaAsnGlnGluProThrProLysSer 1172
 Qy 3822 CCCATGCAATG---GAGAACTGCTTGGCAGCCCCCAGACCGATCGATGACCAACGCC 3878
 Db 1173 ProCysGlyLeuThrGluGlnTyxLeuHisLys-----AspArgTrp---ProGluVal 1189
 Qy 3879 CGCATGACCCCACTATCAAGCCTGCTTCTCACAGAGAGGCTCACGTTTCGCTCCA----- 3932
 Db 1190 SerProGluAspThrGlnSerLeuSerLeuGlu-----SerProSerLys 1206
 Qy 3933 -----CCAGCCGCTCTCAACCCCTGCCACTCTT----- 3959
 Db 1207 GluThrSerLeuAspValSerSerLysGlnLeuSerProGluSerLeuGlyThrLeuGln 1226
 Qy 3960 -----CTGCCCTGAAGAG-----ACTGATGAACCAAGTG 3986
 Db 1227 PheGlyGluLeuAsnLeuGlyLysGluGluMetGlyHisLeuMetGlnAlaGluAspThr 1246
 Qy 3987 ACTCATGATTGCCATCAACTATTGATTGAGGAGACT----- 4022
 Db 1247 SerHisHisThrAlaProMetSerValProGluProHisAlaAlaThrAlaSerProPro 1266
 Qy 4023 -----GGGTCTCCCAAGACCTTACAGACATPACCGCTG----- 4055
 Db 1267 ThrAspGlyThrThrArgTyxSerAlaGlnThrAspIleThrAspAspSerLeuAspArg 1286
 Qy 4056 -----ACTGGAAGAGTCTAACC 4073
 Db 1287 LysSerProAlaSerSerPheSerHisSerThrProSerGlyAsnGlyLysTyxLeuPro 1306
 Qy 4074 TGG-----TTCACTCAGCAAGCAGCTATGTGGTGAA 4106
 Db 1307 GlyAlaIleThrSerProAspGluHisIleLeuThrProAspSerSerPheSerLysSer 1326
 Qy 4107 GGTAAAGAGATGCTGGCGCGGTGGTGAGCGGACCCCGACGATCTGGCCGACGAGC 4166
 Db 1327 ProGluSerLeuProGlyProAlaLeuGluAsp---IleAlaIleLysTrpGluAspLys 1345
 Qy 4167 CTGCCG-----GAAGAACTTTCAGCAAAAGGCTGAGCTCATGGCCCTCAGCGAA 4217
 Db 1346 ValProGlyLeuLysAspArgThrSerGluGlnLysGluProGluProLysAspGlu 1365
 Qy 4218 GCTTTGCGG-----CTGGCCGAAGGGAATCCATA 4247
 Db 1366 ValLeuGlnGlnLysAspLysThrLeuGluHisLysGluValValGluProLysAspThr 1385
 Qy 4248 AACATTTATACGACAGCAGCTATGCTTTTCCGACTGCACACGTA---CATGGGGCCATC 4304
 Db 1386 AlaIleTyxGlnLysAspGlu-----AlaLeuHisValLysAsnGluAlaVal 1401
 Qy 4305 TATAAACAAGGGGTGCTTACTCAGCAGGGAGGGAATAAAGAACAAGAGGAAATTT 4364
 Db 1402 LysGlnGlnAspLysAlaLeuGluGlnLysGlyArgAspLeuGluGlnLysAspThrAla 1421
 Qy 4365 CTAAAGCCTATTAGAAGCCGTACATTTTACCAAAAAGGCTAGCTATTATACACTGCTCTGA 4424
 Db 1422 LeuGluGlnLysAspLysAlaLeuGluPro----- 1431
 Qy 4425 CATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACACAGATGGCTGACCGGGTTGCCAAG 4484
 Db 1432 -----LysAspLysAspLeu-----GluGluLysAspLysAlaLeuGlu 1444

QY	4485	CAGGCAGCCCGGGTGTAACTTCTGCCTATATAGAAATGCCCAAGACCCCGAGACCC	4544
DB	1445	Gln-----LysAspLysIleProGlu---	1451
QY	4545	AGACGACAGTACACCCTAGAAGACTGGCAAGAGATATAAAAGATAGACCACTTCTCTGAG	4604
DB	1452	-----GluLysAspLysAlaLeuGluGln-----	1459
QY	4605	ACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGGAAATCTCTGCCCAACAAGAGGG	4664
DB	1460	-----LysAspThrAlaLeuGluGlnLysAspLysAlaLeuGluProLysAspLysAsp	1477
QY	4665	TTAGAA-----TATGTCCAACAGATACATCGTCTAAACCCACCTAGGAACTAAA	4712
DB	1478	LeuGluGlnLysAspArgValLeuGluGlnLysGluLysIleProGluGluLysAspLys	1497
QY	4713	CACCTGCAGCAGTGTGGTCAGAACATCCCTTTATCATGTCTTCTGAGGCTACCGAGTGGCT	4772
DB	1498	AlaLeuAspGlnLysValArgSer-----	1505
QY	4773	GACTCGGTGGTCAAAACATTTGTGTGCCCTGCCAGCTGGTTATGTCTATCTTCCAGA---	4829
DB	1506	-----ValGluHisLysAlaProGluAspThrValAlaGluMetLysAspArgAsp	1522
QY	4830	--ATGCCTCCAGGGAAGAGACTAAGGGGAAGCACCCAGCGCTCACTGGGAAGTGGAC	4886
DB	1523	LeuGluGlnThrAspLysAlaProGluGlnLysHisGlnAlaGlnGluGlnLysAspLys	1542
QY	4887	TTCACTGAGGTAAAGCCGGCTAAATACGMAAACAAATACCTATTGGTT-----	4934
DB	1543	ValSerGluLysLysAspGlnAlaLeuGluGlnLysTyrTrpAlaLeuGlyGlnLysAsp	1562
QY	4935	-----TTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCTCTAAGAAGAGACT	4988
DB	1563	GluAlaLeuGluGlnAsnIleGlnAlaLeuGluGluAsnHisGlnThrGlnGluGlnGlu	1582
QY	4989	TCAACCGTGGTGTAAAAAAATACTGGAAGAAATTTTTCCAAGATTGGAACTACCTAAG	5048
DB	1583	SerLeuValGlnGluAspLysThrArgLys-----ProLys	1594
QY	5049	GTAATAGGGTCAACAATGGTCCCACTTTTGTGCCAGGTAACTCAGGACTG-----	5102
DB	1595	MetLeu--GluGluLysSerProGluLysValLysAlaMetGluGluLysLeuGluAla	1613
QY	5103	-----GCCAAGATATTGGGATGATTGGAACTGCAATTGTGCATACAGACCC	5150
DB	1614	LeuLeuGluLysThrLysAlaLeuGluLysLeuGluSerLeu-----ValGlnGluGly	1631
QY	5151	CAAAAGCTCAGACAGGTAGAGAGGATGAATAGAACCTTAAGAGACCCCTTACTAAATTG	5210
DB	1632	ArgAlaArgGluGlnGluLysTyrTrpArg-----	1642
QY	5211	ACCGCGAGACTGGCGTTTAATGATTTGGATAGCTCTCTCGCCCTTTGTGCTTTTATGGGTT	5270
DB	1643	--GlyGlnAspValValGlnGluTrpGlnLysThrSerPro-----Thr	1656
QY	5271	AGGAACACCCCTGCACAGTTTGGCTGACCCCTATGAAATTACTTACGGGGGACCCCC	5330
DB	1657	ArgGluGluProAlaGly-----GluGlnLysGluLeu-----Ala	1668
QY	5331	CCATTGGTAGAAATTGCTTCTGTATAGT-----GCTGAC	5366
DB	1669	ProAlaTrpGluAspThrSerProGluGlnAspAsnArgTyrTrpArgGlyArgGluAsp	1688
QY	5367	GTGCTGTTTCCACGCTTTTGTCTTAGCTCAAGGCACCTTAG-----TGGGTG	5417
DB	1689	ValAlaLeuGluGlnAspThrTyrTrpArgGluLeuSerCysGluLysValTrpPhe	1708
QY	5418	AGACAACGAGGTGGAGCAACTCCGGGAGGCCT-----ACTCAGGAGGAGAGACT--	5469
DB	1709	ProHisGluLeuAspGlyGlnGlyAlaArgProHisTyrThrGlnArgGluSerTh	1728

QY	5470	----	-----TG	CAGATCC	CACATC-----	5484	
Db	1728	rPhe	Leu	Asp	Gl	uGlyProAspAspGluGlnGluValProLeuArgGluHisAlaThrAr	1748
QY	5484	-----	-----	-----	-----	5484	
Db	1748	gSer	Pro	Trp	Ala	SerAspPheIysAspPheGlnGluSerSerProGlnIysGlyLeuGl	1768
QY	5485	-GTTT	CCAAGTGGGAGATT	CAGTCTAC	GTGTAGACGCCACCGTGCAGGAAACCTTCGAGACT	5543	
Db	1768	uVal	Gl	uArgTrp	Leu	AlaGluSerProValGlyLeuProProGluGluGluAspIysLe	1788
QY	5544	CGTGTG	GAAGGCCCTTATCTCGTACTTTTGAC	CCACCAACCGCTGTGAAAGTCG	GAAGTCG	AAAGGA	5603
Db	1788	uThr	ArgSer-----	-----	-----	-----	1796
QY	5604	ATCTCC	ACCTCGATCCATCCACGTATAACCGCGCCACCTCCCGATTCCGGGTGG	5663			
Db	1797	-Ser	Pro	Ala	LeuSer-----	ProProGluMetValGlyGlnArgValProSerAlaProGl	1815
QY	5664	AAAGCC	GAAGAACTCAAAATCCCTTAAAGCTTC	TCCATTCACCGCTGCTTCTACTCT			5723
Db	1815	yGln	GluSerProIleProAspProIysLeuMetProHisMet-----	-----	-----	-----	1829
QY	5724	GTC	AAATACCTCTCAGACTAATGTATGCCATAGGACAGCAGCTGA	CTCCATCC	CTCAATAAAC	5783	
Db	1830	-----	-----	-----	-----	-----	1833
QY	5784	CTTAT	CTCTCACCTGGTTAATTACTGACTCCGGC	CACAGGTATTAAATATCAACA	CACTCA	5843	
Db	1833	oThr	ProSerTrpLeu-----	-----	-----	-----	1839
QY	5844	AGGG	AGGCTCTTTTAGGAACCTGGTGGCTGTATCTAC	GTTTGGCTCAGATCAGTTAT	5903		
Db	1840	-----	-----	-----	-----	-----	1853
QY	5904	TCCTAG	CTGACTCACCCTCACCAGATATCTCC	ANGTCCAGGATTTATGTGTTGCC	CAG	5963	
Db	1853	oPro	AlaProLeuSerPro-----	-----	-----	-----	1862
QY	5964	ACCAC	CAAAATATGGAAACAATTCGCGGAATCC	CAGAGATTCTTTTGTATAACA	ATGGAA	6023	
Db	1862	yPro	ProThrProAlaProGluSerHisThrProAlaProPhe-----	-----	-----	-----	1876
QY	6024	CTGTG	TAACTCTAATGATGGATATTGGAAATGGCCAA	CTCTCAGCAGATAGGCTAAAG	6083		
Db	1877	-----	-----	-----	-----	-----	1886
QY	6084	TTTTT	CTTATGTCAACACTATACCAGCTCTGGCAAA	TTTAATTACCTCAGCTGGATTAG	6143		
Db	1887	-----	-----	-----	-----	-----	1891
QY	6144	AACTG	GAAGCCCCAAGTGC-----	-----	-----	-----	6188
Db	1891	nGlu	GlyAlaAlaGluLeuGluGlyGlyProIysSerProLeuGlyIysAspTrpArgly	1911			
QY	6189	AATAA	GTTCCTCAGAAAGGAAACAACAAGAAATATCT	TAAATGGGTAAATGGTATGTC	6248		
Db	1911	g-----	-----	-----	-----	-----	1922
QY	6249	TTGGG	GAATGGTATATATGAGGCTCGGTAAACAACCA	-----	GGCTC	CAATTC	6305
Db	1922	aGlu	AlaProAspIysSerSerHisSerIysValProGluAlaSerLysSerHisAl	1942			
QY	6306	TCGCCT	CAAAATAAACACAGCTGGAGCCTCCATGGCT	-----	-----	-----	6342
Db	1942	aThr	ThrGluProGluGlnThrGluProGluGlnArgGluProThrProIysProAspGl	1962			
QY	6343	-----	-----	-----	-----	-----	6380
Db	1962	uArg	SerPheGlnTyAlaAspIleTyGluGlnMetMetLeuThrGlyLeuGlyProAl	1982			
QY	6381	AA	CCAGGAC	CAGGACCATCTCTAACATAACTCTCGATCAG	ACCCCACTCAG	GTCTAA	6440


```
QY 2412 GCACCCCTCTTAGGTAGAGACTTANTGACCAAGATGGAGACACAAATTTCTTTTAA--- 2468
Db      |||
601 GluGluGlnGlySerLysAspArgGlyLeuAspSerGlyAlaGluThrGluGluGlyLys 620
QY 2469 -----CAAGGGAACACCAAGAGTCTCTCAATAACAAACCTATCATCTGTGTACC 2519
Db      |||
621 AspThrTrpGluGluLysLysGlnArgGluAlaGluArgLeuProAspArgThr----- 638
QY 2520 CTCCAATTAGATGACGAATATCGACTATCTCT---CCCTAGTAAAGCCTGTACAAAT 2576
Db      |||
639 -----GluAlaArgGluGluSerGluProGluValLysGluAspValle 653
QY 2577 ATACAATTCGTTGGACAGTCTCCCAAGCCTGGGCAGAA----- 2618
Db      |||
654 GluLysAlaGluLeuGluGluMetGluGluValHisProSerAspGluGluGluAsp 673
QY 2619 ---ACCGCAGGATGGGTTGGCAAGCAAGTCTCCCAACAGTTATTCACGTGAAGGCC 2675
Db      |||
674 AlaThrLysAlaGluGlyPheTyrGln-----LysHis 684
QY 2676 AGTGCCACACAGTGTCTGACGACAGTACCCCTTGAGTAAAGAGCT-----CAAGAA 2729
Db      |||
685 MetGlnGluProLeuLysValThr-----ProArgSerArgGluAlaPheGlyGly-Ar 702
QY 2730 GGAATTCGGCGCATGTCCAAGATTATCCACAGGGCATCTCTAGTTCTCTGCTCAATCT 2789
Db      |||
702 gGluLeuGlyLeuGlnGlyLysAlaProGluLysGluThrSerLeuPheLeuSerSerLe 722
QY 2790 CCCTGGNAATCTCCCTGCTACCGTTAGAAAGCTCGGACTAATG-----ACTAT 2840
Db      |||
722 u-----ThrThrProAlaGlyAlaThrGluHisValSerTyrIleGlnAspGluThrI 740
QY 2841 CGACCACTACAGCACTTGAGAGAGTCAATAAACGGGTGCAGGATATACACCAACAGTC 2900
Db      |||
740 eProGlyTyrSerGluThrGluGlnThrIleSerAspGluGluLysHisAspGluProGI 760
QY 2901 C-----CGAACCCCTTATAACCTCTTGTGTCTCTCCACCCCAACCGG 2942
Db      |||
760 uGluArgProAlaProProArgPheHisThrSerThrTyrAspLeuProGlyProGluI 780
QY 2943 AGCTGTG-----ATACAGTATTGGACTTAAGGATGCTCTCTCTCCCTGAGA 2990
Db      |||
780 yAlaGlyProPheGluAlaSerGlnProAlaAspSerAlaValProAlaThrSerGlyLy 800
QY 2991 TTAC-----ACCCCACTAGCAACCACTTTTGTGCTTTCGAAT-----GGAGAGAT 3035
Db      |||
800 sValTyrGlyThrProGluThrGluLeuThrTyrProThrAsnIleValAlaAlaProLe 820
QY 3036 CCAGGTACGGGAAGAACCGGCACTCACCTGGA---CCGCACTGCGCCCAAGGTTCAAG 3092
Db      |||
820 uAlaGluGluGluHisValSerAlaThrSerIleThrGluCysAspLysLeuSerSe 840
QY 3093 AACTCCCGCA-----CCATCTTTCAGCAAGCCCTACACAGA 3128
Db      |||
840 rPheAlaThrSerValAlaGluAspGlnSerValAlaSerLeuThrAlaProGlnThrGI 860
QY 3129 GACCTGGCCCAACTTCAGGATCCAACACCTC-----AGGTGACCTC 3170
Db      |||
860 u-----GluThrGlyLysSerLeuLeuAspThrValThrSerIleProSe 877
QY 3171 CTCCAGTACG-----TGATGACCTGCTCTTGGCGGAGGCCACC 3209
Db      |||
877 rSerArgThrGluAlaThrGlnGlyLeuAspTyrVal-ProSerAlaGlyThrIleSerP 897
QY 3210 AAACAGCACTGCTTAGAGGCACGAAGGCACTACTGCTGGAATTGTGACCTAGGCTAC 3269
Db      |||
897 roThrSerSerLeuGluGluAspLysGlyPheLysSer-----ProProCysG 913
QY 3270 AGAGCCTCTGCTAAGAGGCCAGATTTCAGGAGAGAGGTAACTACTTGGGGTACAGT 3329
Db      |||
913 luAspPheSerValThrGly---GluSerGluLysArgGlyGluIleGlyLysGlyL 932
QY 3330 TTGGGAGCGGCGACGCGATGCTGAC----- 3355
Db      |||
932 euSerGlyGluArgAlaValGluGluGluGluThrAlaAsnValGluMetSerG 952
QY 3355 ----- 3355
QY 952 luLysLeuCysSerGlnTyrGlyThrProValPheSerAlaProGlyHisAlaLeuHisP 972
QY 3356 -----GGAGGCACGAGCAAGAAACTGT---AGTCCAGAT----- 3385
Db      |||
972 roGlyGluProAlaLeuGlyGluAlaGluGluArgCysLeuSerProAspAspSerThrV 992
QY 3386 -----ACGGCCCC----- 3394
Db      |||
992 allYsMetAlaSerProProProSerGlyProProSerAlaThrHisThrProPheHisG 1012
QY 3395 -----AACCAACAGCCAAACAATGAGAGAGTTTTTGGGGAGCAGCTGGATTT 3440
Db      |||
1012 lnSerProValGluGluLysSerGluProGlnAspPheGlnGluAlaAspSerTrpGlyA 1032
QY 3441 TGCAGACTGTGATCCCGGGTTTGC-----GACCTTAGCAGCCCCCACTCTACCCGCTA 3494
Db      |||
1032 spThrLysArgThrProGlyValGlyLysGluAspAlaAlaGluThrValLysProG 1052
QY 3495 ACCAAAGAAAAGGGGAATTTCTCC-----TGGGCTCTCTGAG 3530
Db      |||
1052 ly-ProGluGluGlyThrLeuGluLysGluGluLysValProProProArgSerProGln 1071
QY 3531 CACAGAGGATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTC 3590
Db      |||
1072 AlaGlnGluAlaProValAsnIleAspGluGlyLeuThrGlyCysThrIleGlnLeuLeu 1091
QY 3591 CCTGACCTAAACCCCTTTACCTTTATGTGTGATGAG----- 3629
Db      |||
1092 ProAlaGlnAspLysAlaIleValPheGluIleMetGluAlaGlyGluProThrGlyPro 1111
QY 3630 -----CGTAAGGAGTAGCCGGGAGTTTAAACCCAAACCCCTAGGACCATGGAGA 3680
Db      |||
1112 IleLeuGlyAlaGluAlaLeuProGlyGlyLeuArgThrLeuProGlnGluProGlyLys 1131
QY 3681 -----AGACTGTGCTTACCTGTCAAGAAGCTCGATCCTGTAGCCAGTGGT 3728
Db      |||
1132 ProGlnLysAspGluValLeuArgTyrProAspArgSerLeuSerPro----- 1147
QY 3729 TGGCCCATATGCTGAAGGCTATCGCAGCTGTGGCCATCATGTGTCAAGGACGCTGACAAA 3788
Db      |||
1148 -----GluAspAlaGluSer 1152
QY 3789 TTG-----ACTTTGGGACAGAAATATACTGTAAATAGCC 3821
Db      |||
1153 LeuSerValLeuSerValProSerProAspThrAlaAsnGlnGluProThrProLysSer 1172
QY 3822 CCCCATGCAATG---GAGAACATCGTTTCGGCAGCCCCCAGACCGATGATGATGACCAAGCC 3878
Db      |||
1173 ProCysGlyLeuThrGlnTyrLeuHisLys-----AspArgTrp---ProGluVal 1189
QY 3879 CGCATGACCCCATCATCAAGCCCTCTCTCACAGAGAGGGTCAAGTTCGCTCCA----- 3932
Db      |||
1190 SerProGluAspThrGlnSerLeuSerLeuSerGluGlu-----SerProSerLys 1206
QY 3933 -----CCAGCCGCTCTCAACCCCTGCCACTCTT----- 3959
Db      |||
1207 GluThrSerLeuAspValSerSerLysGlnLeuSerProGluSerLeuGlyThrLeuGln 1226
QY 3960 -----CTGCCTGAAGAG-----ACTGATGAACCAAGTG 3966
Db      |||
1227 PheGlyGluLeuAsnLeuGlyLysGluGluMetGlyHisLeuMetGlnAlaGluAspThr 1246
QY 3987 ACTCATGATTGCCATCACTATTATTGATTGAGGAGACT----- 4022
Db      |||
1247 SerHisHisThrAlaProMetSerValProGluProHisAlaAlaThrAlaSerProPro 1266
QY 4023 -----GGGGTCCGCAAGACCTTACAGACATACCGCTG----- 4055
```

Db 1267 ThrAspGlyThrThrArgTyrSerAlaGlnThrAspIleThrAspSerLeuAspArg 1286
::: |||:::||||| |||
Qy 4056 -----ACTGGAGAGTGTCTAAC 4073
|||::: |||
Db 1287 LysSerProAlaSerSerPheSerHisSerThrProSerGlyAsnGlyLysTyrLeuPro 1306
-----TTCACTACAGGAAGCAGCTATGTGGTGAA 4106
TGG----- |||
Qy 4074 1307 GlyAlaIleThrSerProAspGluHisIleLeuThrProAspSerSerPheSerLysSer 1326
::: |||::: |||
Db 4107 GGTAAAGAGTGTGGTGGCGGGTGTGGAGCGGACCGCACGATCTGGCGGCAGCAGC 4166
::: |||::: |||
Db 1327 ProGluSerLeuProGlyProAlaLeuGluAsp---IleAlaIleLysTyrGluAspLys 1345
::: |||::: |||
Qy 4167 CTGCGC-----GAAGAACTTCAGCAAAAGCTGAGCTGATGGCCCTCACGCAA 4217
::: |||::: |||
Db 1346 ValProGlyLeuLysAspArgThrSerGluGlnLysLysGluProGluProLysAspGlu 1365
::: |||::: |||
Qy 4218 GCTTTGCGG-----CTGGCCGGAAGGGAATCCATA 4247
|||::: |||
Db 1366 ValLeuGlnGlnLysAspLysThrLeuGluHisLysGluValValGluProLysAspThr 1385
|||::: |||
Qy 4248 AACATTTATACGACAGCAGGTATGCTTTGCGACTGCACACAGTA---CATGGGGCCATC 4304
|||::: |||
Db 1386 AlaIleTyrGlnLysAspGlu-----AlaLeuHisValLysAsnGluAlaVal 1401
|||::: |||
Qy 4305 TATAACAACAAAGGGGTGCTTACTCTCAGCAGGAGGAAATATAAGAACAAAGAGGAAT 4364
::: |||::: |||
Db 1402 LysGlnGlnAspLysAlaLeuGluGlnLysGlyArgAspLeuGluGlnLysAspThrAla 1421
::: |||::: |||
Qy 4365 CTAAGCCTATTAGAGCGGTACATTACCAAAAAGGCTAGCTATTATACACTGCTCCTGGA 4424
|||::: |||
Db 1422 LeuGluGlnLysAspLysAlaLeuGluPro----- 1431
|||
Qy 4425 CATCAGAAAGCTAAAGATCTCATATCCAGAGGAACACAGATGGCTGACCGGGTTGCCAAG 4484
|||::: |||
Db 1432 -----LysAspLysAspLeu-----GluGluLysAspLysAlaLeuGlu 1444
|||::: |||
Qy 4485 CAGCAGCCAGGGGTGTTAACCTTCTGCTATATAGAAATGCCCAAAGCCCGCAGAACCC 4544
|||::: |||
Db 1445 Gln-----LysAspLysIleProGlu--- 1451
|||
Qy 4545 AGACGACAGTACACCCTAGAAGACTGGCAAGAGATAAAAAGATAGACCGATTCTCTGAG 4604
::: |||::: |||
Db 1452 -----GluLysAspLysAlaLeuGluGln----- 1459
|||
Qy 4605 ACTCGGAAGGACTGCTATACCTCAGATGGGAAGGAATCCTGCCCCCAACAAAGAGAGG 4664
|||::: |||
Db 1460 -----LysAspThrAlaLeuGluGlnLysAspLysAlaLeuGluProLysAspLysAsp 1477
|||
Qy 4665 TTAGAA-----TATGTCCAACAGATACATCGTCTAACCCACCTTAGGAACATAA 4712
|||::: |||
Db 1478 LeuGluGlnLysAspArgValLeuGluGlnLysGluLysIleProGluGluLysAspLys 1497
::: |||::: |||
Qy 4713 CACTGCGAGAGTGTGCTCAGAACATCCCTTATCATGTTCTGAGCTACACGAGGTGCT 4772
|||::: |||
Db 1498 AlaLeuAspGlnLysValArgSer----- 1505
|||
Qy 4773 GACTCGGTGTCAAACTATGTGTGCGCTGCCAGCTGGTTAATGCTTAATCCTTCAGAA--- 4829
|||::: |||
Db 1506 -----ValGluHisLysAlaProGluAspThrValAlaGluMetLysAspArgAsp 1522
|||
Qy 4830 ---ATGCTCCAGGAAGAGACTAAGGGGAAGCCACCGAGCGCTCACTGGGAAGTGGAC 4866
|||
Db 1523 LeuGluGlnThrAspLysAlaProGluGlnLysHisGlnAlaGlnGlnLysAspLys 1542
|||
Qy 4887 TTCAGTAGGTAAAGCCGGCTAAATACGGAACAATACCTATTGGTT----- 4934
::: |||::: |||
Db 1543 ValSerGluLysLysAspGlnAlaLeuGluGlnLysTyrTrpAlaLeuGlyGlnLysAsp 1562
::: |||::: |||
Qy 4935 -----TTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACT 4988
|||::: |||::: |||

Db 1563 GluAlaLeuGluGlnAsnIleGlnAlaLeuGluGluAsnHisGlnThrGlnGluGlnGlu 1582
|||
Qy 4989 TCACCGTGTGTGCTAAAAATACTGGAGAAATTTTCCAAAGATTTGGAAATACCTTAAG 5048
|||::: |||
Db 1583 SerLeuValGlnGluAspLysThrArgLys-----ProLys 1594
|||
Qy 5049 GTATAGGTGCAGACAATGTCACAGCTTTTGTGCCAGGTAAAGTCAAGGAGACTG----- 5102
::: |||::: |||
Db 1595 MetLeu---GluGluLysSerProGluLysValLysAlaMetGluGluLysLeuGluAla 1613
::: |||::: |||
Qy 5103 -----CCCAAGATATTGGGGATTGATGGAAACTGCATTTGTGCATACAGACCC 5150
|||::: |||
Db 1614 LeuLeuGluLysThrLysAlaLeuGlyLeuGluGluSerLeu-----ValGlnGluGly 1631
|||
Qy 5151 CAAAGCTCAGGACAGGTAGAGAGATGAATAGAACCATTAAGAGACCCCTTACTAAATTG 5210
::: |||::: |||
Db 1632 ArgAlaArgGluGlnGluGluLysTyrTrpArg----- 1642
|||
Qy 5211 ACCCGGAGACTGCGCTTAATGATGGATAGTCTCTCTGCCCTTTGTGCTTTTAGGGTT 5270
::: |||::: |||
Db 1643 ---GlyGlnAspValValGlnGluTrpGlnGluThrSerPro-----Thr 1656
|||
Qy 5271 AGAAACACCCCTCGACAGTGGTGGCTGACCCCTATGAATTACTTACGGGGGACCCCCC 5330
|||
Db 1657 ArgGluGluProAlaGly-----GluGlnLysGluLeu-----Ala 1668
|||
Qy 5331 CCATTGGTAGAATGCTCTCTGTACATAGT-----GCTGAC 5366
|||
Db 1669 ProAlaTrpGluAspThrSerProGluGlnAspAsnArgTyrTrpArgGlyArgGluAsp 1688
|||
Qy 5367 GTGCTGTTTCCAGCCTTTGTTCTTAGCTCAAGGCACCTTGAG-----TGGGTG 5417
|||::: |||
Db 1689 ValAlaLeuGluGlnAspThrTyrTrpArgGluLeuSerCysGluArgLysValTrpPhe 1708
|||
Qy 5418 AGAACACGAGCGTGGAGGCAACTCCGGGAGGCT-----ACTCAGAGAGGAGACT-- 5469
::: |||::: |||
Db 1709 -ProHisGluLeuAspGlyGlnGlyAlaArgProHisTyrThrGluGluArgGluSerTh 1728
|||
Qy 5470 -----TGCAATCCACATC----- 5484
|||
Db 1728 rPheLeuAspGluGlyProAspAspGluGlnGluValProLeuArgGluHisAlaThrAr 1748
|||
Qy 5484 ----- 5484
|||
Db 1748 gSerProTrpAlaSerAspPheLysAspPheGlnGluSerSerProGlnLysGlyLeuG 1768
|||
Qy 5485 -GTTTCAAGTGGAGATTGAGTGTAGAGCCGCCCGTGCAGGAGAACTCGAGACT 5543
|||::: |||
Db 1768 uValGluArgTrpLeuAlaGluSerProValGlyLeuProProGluGluAspLysLe 1788
|||
Qy 5544 CGGTGGAGGGCCCTTATCTCGTACTTTTGACCAACACCAACGGCTGTGAAAGTCGAAGGA 5603
|||::: |||
Db 1788 uThrArgSer-----ProPheGluLeuIle 1796
|||
Qy 5604 ATCTCCACCTGGATCCATGCCAGTTTAAACCGGCGCCACCTCCCGATTCGGGGTGG 5663
|||::: |||
Db 1797 -SerProProAlaSer---ProProGluMetValGlyGlnArgValProSerAlaProG 1815
|||
Qy 5664 AAAGCCGAAAAGACTGAAAATCCCTTAAGCTTCGCTCCATCCGCTGGTTCCTTACTCT 5723
|||::: |||
Db 1815 yGlnGluSerProIleProAspProLysLeuMetProHisMet----- 1829
|||
Qy 5724 GTCATAACCTCTCAGACTAATGTTATGCGCATAGGAGACAGCCGTGAACCTCCCATAAACC 5783
|||
Db 1830 -----LysAsnGluPr 1833
|||
Qy 5784 CTTATCTCTCACCTGGTTAATTACTGACTCCGCGCACAGGTATTAATATCAACAACACTCA 5843
|||::: |||
Db 1833 oThrThrProSerTrpLeu----- 1839
|||
Qy 5844 AGGGAGGCTCTTTTAGAACCTGGTGGCTGTATACGTTTGCCTCAGATCAGTTAT 5903
|||::: |||
Db 1840 -----AlaAspIleProProTrpValProLys-----AspArgProLeuPr 1853
|||::: |||

```
QY 5904 TCCTAGTCTGACCTCACCCCCAGATATCCTCCATGCTCACGGATTTTATGTTTCCCCAGG 5963
   |||::: |||||
Db 1853 oProAlaProLeuSerPro-----AlaProGl 1862
QY 5964 ACCACCAAAATATGGAAACATTGCGGAATCCAGAGATTTCTTTTGTAAACAATGGAA 6023
   |||||
Db 1862 yProProThrProAlaProGluSerHisThrProAlaProPhe----- 1876
QY 6024 CTGTGTAACTCTAATGATGATATTGGAATATGGCAACTCTCAGCAGGATAGGGTAAG 6083
   ||| |||::: ||| |||
Db 1877 -----SerTipGlyThrAlaGluTyaspSerVal-- 1886
QY 6084 TTTTCTTANGTCAACACCTATACCAGCTCTGGACAATTTAATTACCTGACCTGGATTAG 6143
   : : : : :
Db 1887 -----ValAlaAlaValGl 1891
QY 6144 AACTGGAAGCCCAAGTGC-----TCTCCTTCAGACCTAGATTACCTAAA 6188
   |||::: ||| |||||
Db 1891 nGluGlyAlaAlaGluLeuGluGlyProTySerProLeuGlyLysAspTyArgLy 1911
QY 6189 AATAAGTTTCACTGAGAAAGAAACAAAGAAATATCCTAAATGGGTAAATGGTATGTC 6248
   | : : : : : : : ||| |||
Db 1911 s-----AlaGluGlyGluArgGluGlu-----GluGlyArgAl 1922
QY 6249 TTGGGGAATGGTATATTATGAGGCTCGGGTAAACAACCA---GGTCCATTCTTAACCTAT 6305
   : ||| ||| ||| |||
Db 1922 aGluAlaProAspLysSerSerHisSerLysValProGluAlaSerLysSerHisAl 1942
QY 6306 TCGCCTCAAAATAAACCCAGCTGGAGCCTCCAAATGGCT----- 6342
   : : : ||| |||||
Db 1942 aThrThrGluProGluThrGluProGluArgGluProThrProTyProAspGl 1962
QY 6343 -----ATAGGACCAATACGGTCTTGACGGGTCAAAGACCCCC 6380
   ||| ||| : : : ||||| |||
Db 1962 uArgSerPheGlnTyAlaAspIleTyGluGlnMetMetLeuThrGlyLeuGlyProAl 1982
QY 6381 AACCCAGGACAGGACCATCTCTTAACATACTTCTGGATCAGACCCCACTGAGTCTAA 6440
   ||| ||| ||| ||| |||
Db 1982 aCysProThrArgGluPro-ProLeuGlyAlaAlaGlyAspTrpProProCysLeuSerT 2002
QY 6441 CAGCAGCACTAAATGGGGCAAACTTTTAGCTCATCCAGGAGCTTTTCAAGCTCT 6500
   ||| ||| : : : : : : : |||
Db 2002 hrLysGluAlaAlaAlaGlyArgAsn---ThrSerAlaGluLysGluLeuSerSerProI 2021
QY 6501 TAACTCCACGA---CTCCAGAGGCTACCTCTTCTTGTGGCTATGCTTAGCTTCGGGCCCC 6557
   : : : : : ||| ||| ||| ||| |||
Db 2021 leSerProLysSerLeuGlnSerAspThrProThrPheSerTyAlaAlaLeu---AlaG 2040
QY 6558 ACCTTACTATGAAGGAATGGCTAGAAGAGGAAATTCATGTGACAAAGACATAGAGA 6617
   ||| : : : ||| ||| |||
Db 2040 lyProThrValProProArgProGluProGlyProSerMet----- 2053
QY 6618 CCAATGCACATGGGATCCCAAAATAGCTTACCTTACTGAGGTTTCTGGAAAAGGCAC 6677
   : : : : : ||| : : : |||
Db 2054 -----GluProSerLeuThrProAlaValPro-----ProArgAlaP 2067
QY 6678 CTGCATAGGAAAGGTTCCCCCATCCCAACACCTTTGTAAACCACCTGAACCTTTAA 6737
   ||| : : : ||| |||
Db 2067 roileLeuSerLysGlyProSerPro----- 2075
QY 6738 TCAAACCTCTGAGAGTCAATATC 6760
   |||||
Db 2076 -----ProLeuAsnGlyAsnIle 2081
```

Search completed: February 14, 2006, 17:20:06
Job time : 220.655 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 14:51:21 ; Search time 9.17949 Seconds
(without alignments)
3295.451 Million cell updates/sec

Title: US-10-723-552-3_COPY_585_2156
Perfect score: 2848
Sequence: 1 ATGGGACAGCGTGACGAC.....CTCTAGAGAAGATAAAGAT 1572

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB spool/US10723552/runat 14022006 125143 12937/app query.fasta_1
-DB=PIR -QFMT=fastan -SUPPIX=rpr -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -HOST=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10723552 @CGN 1 1 107 @runat 14022006 125143 12937 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLONG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRENS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1677.5	58.9	520	1 FOLJGL	gag polyprotein -
2	1637.5	57.5	512	1 FOMVGS	gag polyprotein -
3	1477.5	51.9	537	1 FOMVM7	gag polyprotein -
4	1383.5	48.6	538	2 S35474	gag polyprotein -
5	1381.5	48.5	538	2 S70394	gag polyprotein -
6	1379.5	48.4	538	1 FOMVM	gag polyprotein -
7	1377.5	48.4	1784	2 T10532	gag-pol polyprotei
8	1376.5	48.3	538	1 FOMVIM	gag polyprotein -
9	1371	48.1	537	1 FOMVVR	gag polyprotein -
10	1364	47.9	537	1 FOMVHL	gag polyprotein -
11	1359.5	47.7	540	1 FOMVHL	gag polyprotein -
12	1357	47.6	537	1 FOMVMB	gag polyprotein -
13	1347.5	47.3	536	1 FOMVNE	gag polyprotein -
14	1326	46.6	529	1 FOMVDU	gag polyprotein -

15	1312	46.1	532	1 FOMVVB	gag polyprotein -
16	1239	43.5	536	1 FOMVND	gag polyprotein -
17	1238	43.5	468	1 FOMVMU	gag polyprotein -
18	1049	36.8	427	2 A53798	58k membrane-assoc
19	944	33.1	425	1 FOMVGC	gag polyprotein -
20	930.5	32.7	790	1 FOMVHZ	gag-kit polyprotei
21	778	27.3	371	1 FOMVCS	gag polyprotein -
22	727.5	25.5	608	2 A46312	gag polyprotein -
23	720.5	25.3	310	1 FOMVFB	gag polyprotein -
24	606.5	21.3	559	2 A46311	pol polyprotein -
25	573	20.1	763	1 A40831	gag-akt polyprotei
26	514.5	18.1	313	1 FOMVDA	gag polyprotein -
27	497	17.5	981	1 FOMVGM	gag-abl polyprotei
28	428.5	15.0	187	1 FOMVSS	gag polyprotein -
29	384.5	13.5	663	1 FOMVRR	protein-tyrosine k
30	328.5	11.5	145	2 A32325	gag polyprotein -
31	194.5	6.8	697	2 A26132	gag-abl-pol polypr
32	170.5	6.0	1638	2 A42091	transcription acti
33	165	5.8	1006	2 T42731	atrophin-1 related
34	164	5.8	2526	2 T20531	hypothetical prote
35	164	5.8	2722	2 T20532	hypothetical prote
36	164	5.8	2738	2 E88320	hypothetical prote
37	161	5.7	1184	2 G01763	protein F07A11.6 [
38	160.5	5.6	551	2 S57447	atrophin-1 - human
39	157	5.5	1280	2 T00365	HPBRII-7 protein -
40	154	5.4	618	2 T42664	hypothetical prote
41	154	5.4	707	2 A46302	hypothetical prote
42	150.5	5.3	1263	2 T15496	PTB-associated spl
43	149.5	5.2	887	2 G88484	hypothetical prote
44	149	5.2	701	2 F48613	protein F23F12.8 [
45	149	5.2	701	2 D48613	gag polyprotein -

ALIGNMENTS

RESULT 1

FOLJGL
gag polyprotein - gibbon ape leukemia virus
N;Contains: core protein p10; core protein p12; core protein p15; core protein p30
C;Species: gibbon ape leukemia virus
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: A32595
R;Delassus, S.; Sonigo, P.; Wain-Hobson, S.
Virology 173, 205-213, 1989
A;Title: Genetic organization of gibbon ape leukemia virus.
A;Reference number: A32595; MUID:90051069; PMID:2683360
A;Accession: A32595
A;Molecule type: genomic RNA
A;Residues: 1-520
A;Cross-references: UNIPROT:P21416; UNIPARC:UPI000012AFFF; GB:M26927; NID:g332610; PIDN

Alignment Scores:
Pred. No.: 2.91e-107 Length: 520
Score: 1677.50 Matches: 330
Percent Similarity: 71.1% Conservative: 63
Best Local Similarity: 59.7% Mismatches: 93
Query Match: 58.9% Indels: 67
DB: 1 Gaps: 10

US-10-723-552-3_COPY_585_2156 (1-1572) x FOLJGL (1-520)

Qy 1 ATGGGACAGCGTGACGACCCCTCTAGTTGACTCTCGACCTTGGACTGAAGTTAAA 60
Db 1 MetGlycInAspAsnSerThrProfilSerLeuThrLeuAsnHisTrpArgAspValArg 20


```
QY 61 TCCAGGGCTCATATTTCTCAGTTCAGTTAAGAGGACCTTGGCAGACTTTCGTGTC 120
Db 21 ThrArgAlaHisAsnLeuSerValLysIleArgLysGluYsTrpGlnThrPheCysSer 40
QY 121 TCTGAATCGCCGACATTCGATCGATGGCCATCAGAGGGACCTTTTAATTCAGATT 180
Db 41 SerGluTrpProThrPheGlyValGlyTrpProProGluGlyThrPheAsnLeuSerVal 60
QY 181 ATCTGGCTGTGTTAAAGCAGTTATTTTTCAGACTGGACCCGGCTCTCATCCGATCAGGAG 240
Db 61 IlePheAlaValLysArgIleValPheGlnGlu---ThrGlyGlyHisProAspGlnVal 79
QY 241 CCCTATATCTTACCTGCGCAGATTTGCGAGGATCTCTCCGCATGGGTTAAACCATGG 300
Db 80 ProTrpIleValValTrpGlnAspLeuAlaGlnSerProProTrpValProProSer 99
QY 301 CTGAATAGCCCAAGAAAGCCAGGTCCTCCGAAATTCCTGGCTCTTGGAGAGAAAAACAACAC 360
Db 100 AlalysIleAlaValValSerSerProGluAsnThrGlnGlyProSerAlaGlyArgPro 119
QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAG----- 402
Db 120 SerAla-----ProProArgProProIleTyProAlaThrAspAspLeuLeu 136
QY 403 ATTGAGGACCCACCGGCTTGGCGGAACCCCAATCTGTTCGCCACCCCTTATCTGGCA 462
Db 137 LeuSerGluProPro-----ProTyProAlaAlaLeuProProPro-----LeuAla 152
QY 463 CAGGTCGCCGGAGGGACCTTTGCCCTCTCGAGCTCCG-----GCCGTGGAG 513
Db 153 ProProAlaVal----GlyPro---AlaProGlyGlnAlaProAspSerSerAspProGlu 170
QY 514 GGACTGTGTCAGGAGCTCGAGCGCGGAGCGGCCACCCCG-----GAGCGGACA 564
Db 171 GlyProAlaAlaGlyThrArgSerArgArgAlaArgSerProAlaAspAspSerGlyPro 190
QY 565 GACGAGATCGGCAATTACCTCGGCAGCTGACGGCCCTCCACACCGGGGGGCCAATG 624
Db 191 AspSerThrValIleLeuProLeuArgAlaIleGlyProProAlaGluProAsnGlyLeu 210
QY 625 CAGCCCTCCAGTATTGGCCCTTTCTTCGCAGATCTCTATATTTGGAACCTAACCAT 684
Db 211 ValProLeuGlnTyTrpProPheSerSerAlaAspLeuTyAsnTrpLysSerAsnHis 230
QY 685 CCCCTTTCTCGGAGGATCCCAACGCTCACGGGGTTGGTGAGTCCCTTATCTCTCT 744
Db 231 ProSerPheSerGluAsnProAlaGlyLeuThrGlyLeuLeuGluSerLeuMetPheSer 250
QY 745 CACAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTTCACAAACGAGGAG 804
Db 251 HisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnIleLeuPheThrThrGluGlu 270
QY 805 CGAGAGAGATTTCTATTAGAGGCTAGAAAAATGTTCTCTGGGCGGACGGCGACCCACG 864
Db 271 ArgGluArgIleLeuLeuGluAlaArgLysAsnValLeuGlyAspAsnGlyAlaProThr 290
QY 865 CGGTTCAAAATGAGATTGACATGGGATTTCCCTTAACTCGCCCGGTTGGGACTACAAC 924
Db 291 GlnLeuGluAsnLeuIleAsnGluAlaPheProLeuAsnArgProGlnTrpAspTyAsn 310
QY 925 ACGCTGAAGTAGGGAGAGCTTGAATAATCTATCGCAGGCTCTGTGGCGGGTCTCCGG 984
Db 311 ThrAlaAlaGlyArgGluLeuLeuValTyArgArgThrLeuValAlaGlyLeuLys 330
QY 985 GCGCCTCAGACGGCCCATTAATTTGGCTTAAGTAAAGAGATGATGCGAGGACCGAAT 1044
Db 331 GlyAlaAlaArgArgProThrAsnLeuAlaLysValArgGluValLeuGlnGlyProAla 350
QY 1045 GAACCCCTCTGTTTTCTTGAGGCTCTTGGAGCCCTTCAGGGCGGTACACCCCTTTT 1104
Db 351 GluProProSerValPheLeuGluArgLeuMetGluAlaTyArgArgTyThrProPhe 370
QY 1105 GATCCCACCTCAGAGGGCCCAAAAGCCTCAGTGGCTTTTGGCCCTTTATAGGACAGTCAGCC 1164
```

```
Db 371 AspProSerGluGluGlyGlnGlnAlaAlaValAlaThrAlaPheThrGlyGlnSerAla 390
QY 1165 TTGATATTAGAAAGAGCTTCAGAGACTGGAAGGTTTACAGGAGGCTGAGTTACGTGAT 1224
Db 391 ProAspIleLysLysLysLeuGlnArgLeuGluGlyLeuGlnAspTySerLeuGlnAsp 410
QY 1225 CTAGTGAAGGAGGAGAGAAAGTATATTACAAAAGGAGACAGAGAAGAAAGGACAA 1284
Db 411 LeuValArgGluAlaGluLysValTyHisLysArgGluThrGluGluArgGlnGlu 430
QY 1285 AGAAAAGAGAGAGAGAGAGAGAAAGGAGAGAAAGCTAATAACCGCAAGAGAGAAAT 1344
Db 431 ArgGluLysLysGluAlaGluGluArgGluArgArgAspArgArgGlnGluLysAsn 450
QY 1345 TTGACTAAGATCTTGGCTGCGAGTGGTT---GAAGGAAAGCAATACCGAAAGAGAGA 1401
Db 451 LeuThrArgIleLeuAlaAlaValValSerGly----- 462
QY 1402 GATTTTAGGAAATTAGGTGAGCCCTAGACAGTCAGGAACTGGGCAATAGGACC--- 1458
Db 463 -----GlySerGlyAspArgArgThrGlyAsnLeuGlyAsnArgAlaArg 477
QY 1459 -----CCACTCGACAAGGACCAATGTGCATATTGTAAAGAA 1494
Db 478 LysThrProArgAspGlyArgProProLeuAspLysAspGlnCysAlaTyCysLysGlu 497
QY 1495 AGAGGACACTGG 1506
Db 498 LysGlyHisTrp 501

RESULT 3
FOMVM7
gag polyprotein - baboon endogenous virus (strain M7)
N;Contains: core protein p15; core shell protein p30; inner coat protein p12; nucleoprotein p12;
C;Species: baboon endogenous virus
A;Note: host Papio sp. (baboon)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: J02060
R;Kato, S.; Matsuo, K.; Nishimura, N.; Takahashi, N.; Takano, T.
Jpn. J. Genet. 62, 127-137, 1987
A;Title: The entire nucleotide sequence of baboon endogenous virus DNA: a chimeric gene
A;Reference number: J02060
A;Accession: J02060
A;Molecule type: DNA
A;Residues: 1-537 <KAT>
A;Cross-references: UNIPROT:P03341; UNIPARC:UPI000012AFEE; GB:D10032; GB:D00088; NID:g2
C;Comment: This protein is synthesized as a gag-pol polyprotein.
C;Genetics:
A;Gene: gag
C;Superfamily: mammalian retrovirus gag polyprotein 1
C;Keywords: core protein; inner coat protein; nucleoprotein; polyprotein
F;1-112/Product: inner coat protein p12 #status predicted <INP>
F;113-226/Product: core protein p15 #status predicted <CPP>
F;227-477/Product: core shell protein p30 #status predicted <CSP>
F;478-537/Product: nucleoprotein p10 #status predicted <NPP>

Alignment Scores:
Pred. No.: 1,49e-93 Length: 537
Score: 1477.50 Matches: 293
Percent Similarity: 66.4% Conservative: 80
Best Local Similarity: 52.1% Mismatches: 126
Query Match: 51.9% Indels: 63
DB: 1 Gaps: 10

US-10-723-552-3_copy_585_2156 (1-1572) x FOMVM7 (1-537)
QY 1 ATGGGACAGAGCGTGACGACCCCTCTTAGTTTCTCACTCTCGACCATTTGGAGCTGAAGTTAAA 60
Db 1 MetGlyGlnThrLeuThrThrProLeuSerLeuThrLeuThrHisPheSerAspValArg 20
QY 61 TCCAGGCGCTCATATTTGTGTCAGTTAAGAGGACCTTGGCAGACTTTCGTGTC 120
:::|||||
```

```
Db      21  AlaArgAlaHisAsnLeuSerValGlyValArgLysGlyArgTrpGlnThrPheCysSer  40
Qy      121 TCTGAATGCCCGCATTCGATGTTGGATGGCCATCAGAGGGAGCTTTAAATTCGAGATT 180
Db      41  SerGluTrpProThrLeuHisValGlyTrpProArgAspGlyThrPheAspLeuSerVal  60
Qy      181 ATCTCGGCTGTAAAGCAGATTATTTTTCAGACTCGACCCGGCTCTCATCCCGATCAGGAG 240
Db      61  IleLeuGlnValLysThrLysValMetAspProGlyProHisGlyHisProAspGlnVal  80
Qy      241 CCTATATCTTACGTGGCAAGATTTCGACAGGATCTCCGCCATGGTGTAAACCATGG 300
Db      81  AlatyrlleIleThrTrpGluAspLeuValArgAsnProProProTrpValLysProPhe  100
Qy      301 CTGAATAGCCAGAAAGCCAGGTCCCGAATTCGGCTCTTCGAGAGAAACAAACAC 360
Db      101 LeuHisThrProSerThrSerLysSerThrLeuLeuAlaLeuGluValProLysAsnArg  120
Qy      361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAG----- 402
Db      121 ThrLeuAspProProLysPro-----ValLeuProAspGluSerGlnGlnAspLeu  137
Qy      403 -----ATTGAGGAGCCACCG-----GCTGGCCCGGAACCCCAA 435
Db      138 LeuPheGlnAspProLeuProHisProProHisAsnProLeuLeuGluProProProTyr  157
Qy      436 TCTGTTCCCCCACCCTTATCTGGCACAGGTGCCGGAGGGACCTTTGCC----- 489
Db      158 AsnSerProSerProProValLeuSerProValSerProThrThrProSerAlaProThr  177
Qy      490 -----CCTCCTGGAGCTCCGGCG----- 507
Db      178 ProSerSerLeuValSerSerThrProProSerSerProAlaProProGluLeuThr  197
Qy      508 -----GTGGAGGACCTCTCGACGGGACTCGGAGCGCGAGGGCGCCACCCCG 555
Db      198 ProArgThrProProGlnThrProArgLeuArgLeuArgAlaGluGlyGlnAspGly  217
Qy      556 GAGCGGACAGACGAGATCGACATTAACGCTGGCGCAGGTACGGCCCTCCACACCGGG 615
Db      218 ProSerThrTrpGlnSerSerLeuPheProLeuArgThrValAsn----- 232
Qy      616 GCCCAATTGACGCCCTCCAGATTATGGCCCTTCTCTCGAGATCTCTATAATTGGAAA 675
Db      233 -----ArgThrIleGlnIleTrpProPheSerAlaSerAspLeuTyrAsnTrpLys  249
Qy      676 ACTAACCATCCCTTCTCGAGAGATCCCAACGCCCTCAGGGGTTGGTGAGTCCCTT 735
Db      250 ThrHisAsnProSerPheSerGlnAspProGlnAlaLeuThrSerLeuIleGluSerIle  269
Qy      736 ATGTTCTCTACACGCTTCTGGATGATTGTCAACAGCTGCTCGACACACTCTTTCACA 795
Db      270 LeuLeuThrHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnValLeuLeuThr  289
Qy      796 ACCGAGGAGCGAGACAGATTCTATTAGAGGCTAGAAAAATGTTCTCGGGCGCGACGG 855
Db      290 ThrGluGluArgGlnArgLeuLeuLeuGluAlaArgLysAsnValProGlyProGlyGly  309
Qy      856 GCACCCAGCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTAACTCGCCCGCGTTGG 915
Db      310 LeuProThrGlnLeuProAsnGluIleAspGluGlyPheProLeuThrArgProAspTrp  329
Qy      916 GACTACACACGGCTGAAGGTAGGAGAGCTTGAAATCTATCGCCAGGCTCTGTGGCG 975
Db      330 AspTyrGluThrAlaProGlyArgGluSerLeuArgIleTyrArgGlnAlaLeuLeuAla  349
Qy      976 GTCCTCCGGGCGCTCAAGACGGCCCTAATTTGGCTTAAGTAAAGAGAGTGTGACG 1035
Db      350 GlyLeuLysGlyAlaGlyLysArgProThrAsnLeuAlaLysValArgThrIleThrGln  369
Qy      1036 GGACCGAATGAACCCCTCTCTTTTCTTAGAGGCTCTTGGAAAGCCCTTCAGCGGGTAC 1095
Db      370 GlyLysAspGluSerProAlaAlaPheMetGluArgLeuLeuGluGlyPheArgMetTyr  389
```

```
Qy      1096 ACCCTTTTGATCCCACTCAGAGGCCCAAAAAGCCTCAGTGGCTTTGGCCTTTATAGGA 1155
Db      390 ThrProPheAspProGluAlaProGluHisLysAlaThrValAlaMetSerPheIleAsp  409
Qy      1156 CAGTCAGCCTTGGATATTGAAGAAGCTTCAGAGACTGGAAGGTTCAGGAGGCTGAG 1215
Db      410 GlnAlaAlaLeuAspIleLysGlyLysLeuGlnArgLeuAspGlyIleGlnThrHisGly  429
Qy      1216 TTACCTCATCTAGTGAAGGAGGAGAAAGTATATTACAAAAGGGAGACAGAAAGAA 1275
Db      430 LeuGlnGluLeuValArgGluAlaGluLysValTyrAsnLysArgGluThrProGluGlu  449
Qy      1276 AGGGAACAAAGAAAGAGAGAAAGAGAAAGAGAAAGGAGGAGAAACCTTAATAACGGCAA 1335
Db      450 ArgGluAlaArgLeuIleLysGluGlnGluArgGluAspArgAspArgLysArg  469
Qy      1336 GAGAAGAATTGACTAAGATCTTGGCTGCAAGTGTGAAAGGAAAGCAATACCGAAAGA 1395
Db      470 AspLysHisLeuThrLysIleLeuAlaAlaValVal-----ThrGluLys  484
Qy      1396 GAGAGAGATTTAGGAAAATTAGTTCAGGCCCTAGACAGTCAGGGAACCTGGGCAATAGG 1455
Db      485 -----ArgAlaGly-----LysSerGlyGluThrArgArgArg  495
Qy      1456 ACCCCACTCGACAGGACCAATGTGCATATTGTAAGAAAGAGGACACTGGCGCAAGCAAC 1515
Db      496 ProLysValAspLysAspGlnCysAlaTyrCysLysGluArgGlyHisTrpIleLysAsp  515
Qy      1516 TGCCCCCAAGAAAG-----GGAAACAAAGGACCAAGGATCTTAGCTCTTAGAAGAAGAT 1566
Db      516 CysProLysArgProArgAspGlnLysProAlaProValLeuThrLeuGlyGluAsp  535
Qy      1567 AAAGAT 1572
Db      536 SerGlu 537

RESULT 4
R35474
gag polyprotein - Friend murine leukemia virus (strain PVC-211)
N;Contains: core protein p15; core shell protein p30; inner coat protein p12; nucleoprotein p10;
C;Species: Friend murine leukemia virus
A;Variety: strain PVC-211
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
R;Remington, M.P.; Hoffman, P.M.; Ruscetti, S.K.; Masuda, M.
Nucleic Acids Res. 20, 3249, 1992
A;Title: Complete nucleotide sequence of a neuropathogenic variant of Friend murine leukemia virus
A;Reference number: S35474; MUID:92319660; PMID:1620621
A;Accession: S35474
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-538 <REM>
A;Cross-references: UNIPROT:P26805; UNIPARC:UPI000012B02C; EMBL:M93134; NID:9331898; PID:
A;Experimental source: strain PVC-211
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992
C;Genetics:
A;Gene: gag
C;Superfamily: mammalian retrovirus gag polyprotein I
C;Keywords: core protein; inner coat protein; nucleoprotein; polyprotein

Alignment Scores:
Pred. No.: 4.14e-87 Length: 538
Score: 1383.50 Matches: 285
Percent Similarity: 62.3% Conservative: 69
Best Local Similarity: 50.2% Mismatches: 137
Query Match: 48.6% Indels: 77
DB: 2 Gaps: 14

US-10-723-552-3_COPY_585_2156 (1-1572) x S35474 (1-538)

Qy      1  ATGGACGACGCGTGTACGACCCTCTTAGTTTGACTCTCGACCATTCGACCTGAAGTTAAA 60
Db      1  ATGGACGACGCGTGTACGACCCTCTTAGTTTGACTCTCGACCATTCGACCTGAAGTTAAA 60
```

Db 1 MetGlyGlnThrAlaThrThrProLeuSerLeuThrLeuAspHisTrpLysAspValGlu 20
Qy 61 TCCAGGGCTCAATAATTGTCAGTTCAGGTTAAGAGGACCTTGGCAGACTTCTGTGTC 120
Db 21 ArgThrAlaHisAsnGlnSerValGluValArgLysArgArgTrpValThrPheCysSer 40
Qy 121 TCTGAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAATCTGAGATT 180
Db 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnProAspIle 60
Qy 181 ATCTCTGGCTGTTAAAGCAGTATTATTTTCAGACTGACCCGCTCTCATCCGATCAGGAG 240
Db 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
Qy 241 CCTATATCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATCGGTTAAACCATGG 300
Db 81 ProFyrIleValThrTrpGluAlaLeuAlaValAspProProTrpValLysProPhe 100
Qy 301 CTGAATAAGCCAAAGACCGAGTCCCGGAATTCCTGGCTCTTGGAGAGAAAAACAACAC 360
Db 101 ValHis-----ProLysProProLeuLeu-----ProPro 111
Qy 361 TCGGCTGAAAAGTCAAGCCCTCTCTCT-----CATATCTAC 396
Db 112 SerAlaProSerLeuProGluProProLeuSerThrProGlnSerSerLeuTyr 131
Qy 397 CCCGAGATTGAGGACCA---CCGGCTTGGCCGGAACCCCAATCTGTTCCTC----- 444
Db 132 ProAlaLeuThrSerProLeuAsnThrLysProArgProGlnValLeuProAspSerGly 151
Qy 445 -----CCACCCCTTATCTGGCACAAGGT----- 468
Db 152 GlyProLeuIleAspLeuThrGluAspProProProTyrArgAspProGlyProPro 171
Qy 469 -----GCCGCGAGGGACCTTTCCTCTC-----CCTGAGCTCCGCGGTG 510
Db 172 SerProAspGlyLysGlyAspSerSerGlyGluValAlaProThrGluGlyAlaProAspSer 191
Qy 511 GAGGGACCTGCTGAGGAGTCCGAGCGCGGAGGGCGCCACCCCGAGCGACAGACGAG 570
Db 192 SerProMetValSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrThr 211
Qy 571 ATCGCGACATTACGCTGCGCACGTACGGCCCTCCACACCGGGGGGCCAAATTCAGCCCC 630
Db 212 SerGlnAlaPheProLeuArgLeuGlyGly-----AsnGlyGln----- 224
Qy 631 CTCAGTATTGGCCCTTTCTTCGACAGTCTATATAATTGGAAAACTAACCATCCCTCT 690
Db 225 PheGlnTrpProPheSerSerSerAspLeuTyrAsnTrpLysAsnAsnAsnProSer 244
Qy 691 TTCTCGGAGGATCCCAACGCTCACCGGGTGTGGAGTCCCTTATGTTCTCTCACCAG 750
Db 245 PheSerGluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuLeuThrHisGln 264
Qy 751 CCTACTTGGGATGATTCTCAACAGTCTGTCAGACACTCTTCAACACCGAGGACGAGAG 810
Db 265 ProThrTrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGln 284
Qy 811 AGAATTCATTAGAGGTAGAAAAAATGTTCTCGGGCCGACCGGGACCCACGCGGTG 870
Db 285 ArgValLeuLeuGluAlaArgLysAlaValArgGlyGluAspGlyArgProThrGlnLeu 304
Qy 871 CAATAATGAGATTGACATGGATTTCCTTAACCTCGCCCGGTGGGACTACAACAGCGT 930
Db 305 ProAsnAspIleAsnAspAlaPheProLeuGluArgProAspTrpAspTyrAsnThrGln 324
Qy 931 GAAGGTAGGAGAGTGAATAATCTATCGCCAGGCTCTGGTGGCGGTCTCCGGGGGCC 990
Db 325 ArgGlyArgAsnHisLeuValHisTyrArgGlnLeuLeuAlaGlyLeuGlnAsnAla 344
Qy 991 TCAAGACGGCCCACTAATTTGGCTTAAGTAAGTAAGTATGACGAGGACCGAATGACCC 1050
Db 345 GlyArgSerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSer 364

Qy 1051 CCCTCTGTTTTTTCTTGAGAGCTCTTGAAGCCTTTCAGGGCGGTACACCCCTTTTATCCCC 1110
Db 365 ProSerAlaPheLeuLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspPro 384
Qy 1111 ACCTCAGAGCCCAAAAAGCCTCAGTGGCTTGGCTTTATAGGACAGTACGCTTGAT 1170
Db 385 GluAspProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAsp 404
Qy 1171 ATTAGAAGAGCTTTCAGAGACTCGAAGGGTTACAGGAGGCTGAGTTACGTGATCTAGTG 1230
Db 405 IleGlyArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuVal 424
Qy 1231 AAGGAGCGCAGAAAGTATATTACAAAAGGGAGAGACAGAAGAAAGGAAACAAAGAAA 1290
Db 425 ArgGluAlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgVal 444
Qy 1291 GAGGAGAAAGAGAGAAAGGAGGAAAGA----- 1320
Db 445 ArgArgGluThrGluGluLysGluGluArgArgAlaGluAspGluArgGluLys 464
Qy 1321 ---CGTAATAAAGCGCAAGAGAGAAATTTGACTTAAGATCTTGGCTGAGTGGTTGAGGG 1377
Db 465 GluArgAspArgArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGly 484
Qy 1378 AAAAGCAATACGGAAAGAGAGAGAGATTTTAGGAAATTAGGTCCAGCCCTAGACAGTCA 1437
Db 485 Gln-----ArgGlnAspArgGlnGlyGlyGluArg--- 494
Qy 1438 GGGAACTGGGCAATAGGACCCCTCGACAAGGACCAATGTGCATATTGTAAGAAAGA 1497
Db 495 -----ArgArgProGlnLeuAspHisAspGlnCysAlaTyrCysLysGluLys 510
Qy 1498 GGCACCTGGCAAGGAACTGCCCCAGAG-----GGAAACAAGGACCAAGG----- 1545
Db 511 GlyHisTrpAlaArgAspCysProLysLysProArgGlyProArgGlyProArgProGln 530
Qy 1546 -----ATCCTAGCTCTAGAAGAA 1563
Db 531 AlaSerLeuLeuThrLeuAspAsp 538

RESULT 5

S70394 gag polyprotein - Friend murine leukemia virus (strain FB29)

N:Contains: core protein p15; core shell protein p30; inner coat protein p12; nucleopro

C:Species: Friend murine leukemia virus

A:Variety: strain FB29

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S70394

R;Perryman, S.; Nishio, J.; Chesebro, B.

Nucleic Acids Res. 19, 6950, 1991

A:Title: Complete nucleotide sequence of Friend murine leukemia virus, strain FB29.

A:Reference number: S70393; MUID:92107687; PMID:1762923

A:Accession: S70394

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-538 <PER>

A:Cross-references: UNIPROT:P26806; UNIPARC:UP1000012802B; EMBL:Z11128; NID:G61547; PID

A:Experimental source: strain FB29

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991

C:Genetics:

A:Gene: gag

C:Superfamily: mammalian retrovirus gag polyprotein I

C:Keywords: core protein; inner coat protein; nucleoprotein; polyprotein

Alignment Scores:

Pred. No.:	5,67e-87	Length:	538
Score:	1381.50	Matches:	285
Percent Similarity:	62.9%	Conservative:	72
Best Local Similarity:	50.2%	Mismatches:	134
Query Match:	48.5%	Indels:	77
DB:	2	Gaps:	15

C;Keywords: polypeptide
F:2-131/Product: core protein p15 #status predicted <P15>
F:132-215/Product: inner coat protein p12 #status predicted <P12>
F:216-478/Product: core shell protein p30 #status predicted <P30>
F:479-534/Product: nucleoprotein p10 #status predicted <P10>

Alignment Scores:
Pred. NO.: 7,78e-87 Length: 538
Score: 1379.50 Matches: 280
Percent Similarity: 61.8% Conservative: 77
Best Local Similarity: 48.4% Mismatches: 124
Query Match: 48.4% Indels: 97
DB: 1 Gaps: 13

US-10-723-552-3_COPY_585_2156 (1-1572) x FOMVM (1-538)

QY 1 ATGGGACAGCGGTGACGACCCCTCTTAGTTTGAATCTCGACCATTTGAGTAAAT 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpLysAspValGlu 20
QY 61 TCCAGGGCTCATAATTTGTCAGTTTTCAGGTTAAGAGGACCTTGGCAGACTTTCTGTGTC 120
DB 21 ArgLeuAlaHisAenGlnSerValAspValLysLysArgArgTprValThrPheCysSer 40
QY 121 TCTGAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATCTCGATT 180
DB 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
QY 181 ATCTGGCTGTAAAGCAGTTATTTTCAGACTGACCGGCTCTCATCCCGATCAGAG 240
DB 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
QY 241 CCTTATATCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATCGGTTAAACCATGG 300
DB 81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProProTrpValLysProPhe 100
QY 301 CTGAATTAAGCAAGAACCGAGTCCCGAATTCCTGGCTCTTGGAGAGAGAAAAACAACAC 360
DB 101 ValHis---ProLysProProProProLeuLeuProSer----- 112
QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCATATCTACCCGAGATTGAGAGCCACCG--- 417
DB 113 -----AlaProSerLeuProLeuGluProProLeuSerThrProProGln 127
QY 418 -----GCTTGGCCGGGAACCCCAA----- 435
DB 128 SerSerLeuTyrProAlaLeuThrProSerLeuGlyAlaLysProLysProGlnValLeu 147
QY 436 -----TCTGTTCCCCCACCCTTATCTGGCA 462
DB 148 SerAspSerGlyProLeuIleAspLeuLeuThrGluAspProProProTyr----- 165
QY 463 CAGGGTCCCGGAGGGGACCTTTGGCCCTCCT----- 495
DB 166 -----ArgAspProArgProProProSerAspArgAspGlyAspSerGlyGlu 181
QY 496 GGAGCTCCGGCGGTGGAGGACCT-----GCTCAGGAGACTCGGAGCCGG 540
DB 182 AlaThrProAlaGlyGluAlaProAspProSerProMetAlaSerArgLeuArgGlyArg 201
QY 541 AGGGGCCACCCCGGAGCGGACAGAGATCCCGCATTCACGCTGGCAGCATCGGC 600
DB 202 ArgGluProProValAlaAspSerThrThrSerGlnAlaPheProLeuArgThrGlyGly 221
QY 601 CCTCCCACACGGGGGGCCATTCGACGCCCTCCAGTATTGGCCCTTTCTTCTGCAGAT 660
DB 222 -----AenGlyGln-----LeuGlnTyrTrpProPheSerSerAsp 234
QY 661 CTCTATAATTTGAAAACTAACCATTCCTTCTCGAGGATCCCAACGCTCAGCGGG 720
DB 235 LeuTyrAsnTrpLysAsnAsnAsnProSerPheSerGluAspProGlyLysLeuAla 254
QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACCAGCCTACTTGGGATGATTGTCAACAGCTGCTG 780

DB 255 LeuIleGluSerValLeuIleThrHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 274
QY 781 CAGACACTTCTTACACCGGAGCGGAGAGAAATTCATTAGAGGCTAGAAAAAATGTT 840
DB 275 GlyThrLeuLeuThrGlyGluGlyGlnArgValLeuLeuGluAlaArgLysAlaVal 294
QY 841 CTGGGGCCGACGGCGGACCCACCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTA 900
DB 295 ArgGlyAspAspGlyArgProThrGlnLeuProAsnGluValAspAlaPheProLeu 314
QY 901 ACTCGCCCGGTTGGGACTCAACACCGCTCAAGCTAGGAGAGCTTGAATCTATCGC 960
DB 315 GluArgProAspTrpGluTyrThrGlnAlaGlyArgAsnHisLeuValHisTyrArg 334
QY 961 CAGGCTCTGGTGGCGGCTCTCCGGCGCCTCAAGACGGCCCACTTAATTCGCTAAGTA 1020
DB 335 GlnLeuLeuIleAlaGlyLeuGlnAsnAlaGlyArgSerProThrAsnLeuAlaLysVal 354
QY 1021 AGAAGTGTATGACGAGGACCGAATGAACCCCTCTGTTTTCTTGAGAGGCTCTTGAA 1080
DB 355 LysGlyIleThrGlnGlyProAsnGluSerProSerAlaPheLeuGluArgLeuLysGlu 374
QY 1081 GCCTTCAGGCGGTACACCCCTTTGATCCACCTCAGAGGCCCAAAAGCCCTCAGTGCT 1140
DB 375 AlaTyrArgArgTyrThrProTyrAspProGluAspProGlyGlnGluThrAsnValSer 394
QY 1141 TTGGCTTTTATAGACAGTCTGATATTTAGAAAGAGCTTCAGAGACTCGAAGG 1200
DB 395 MetSerPheIleTrpGlnSerAlaProAspIleGlyArgLysLeuGluArgLeuLys 414
QY 1201 TTACAGAGGCTGAGTTACGTGATCTAGTGAAGAGGACGACAGAAAGTATATTACAAAG 1260
DB 415 LeuArgAsnLysThrLeuGlyAspLeuValArgGluAlaGluArgIlePheAsnLysArg 434
QY 1261 GAGCACAAGAAGAAGGGAACAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 435 GluThrProGluArgGluArgGluArgIleArgArgGluArgGluArgLysGluArg 454
QY 1321 -----CGTAAACCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1347
DB 455 ArgArgThrGluAspGluGlnLysGluArgAspArgArgGHisArgGluMet 474
QY 1348 ACTAAGATCTTGGCTGAGTGGTGAAGGGAAGCAATACGGAAGAGAGAGAGAGAGATTT 1407
DB 475 SerArgLeuLeuAlaThrValValSerGlyGlnArgGlnAspArgGlnGluArg 494
QY 1408 AGGAAATTAGTTCAGGCCCTCAGACATCAGGAGAACTGGGCAATAGAGACCCCACTGAC 1467
DB 495 Arg-----ArgSerGlnLeuAsp 500
QY 1468 AAGGACCAATGTCATATTGTAAAGAGAGGACACTGGGCAAGAACTGCCCAAG--- 1524
DB 501 CysAspGlnCysThrTyrCysGluGluGlnGlnHisTrpAlaLysAspCysProArg 520
QY 1525 ---AAGGAAACAAAGGACCAAGG-----ATCCTAGCTCTAGAGAA 1563
DB 521 ProArgGlyProArgGlyProArgProGlnThrSerLeuLeuThrLeuAspAsp 538

RESULT 7

T10532

gag-pol polyprotein precursor - feline leukemia virus (strain FeLV-FAIDS)

N;Contains: core protein p10; core protein p12; core protein p15; core protein p27; end

C;Species: feline leukemia virus

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: T10532

R;Donahue, P.R.; Hoover, E.A.; Beltz, G.A.; Riedel, N.; Hirsch, V.M.; Overbaugh, J.; Mu

J. Virol. 62, 722-731, 1988

A;Title: Strong sequence conservation among horizontally transmissible, minimally patho

A;Reference number: Z17078; MUID:88119207; PMID:2828667

A;Accession: T10532

A;Status: translated from GB/EMBL/DDJ

A;Molecule type: mRNA

A:Title: Nucleotide sequence of Moloney murine leukaemia virus.

A:Reference number: A93265; MUID:82035843; PMID:6169994

A:Accession: A03930

A:Molecule type: Genomic RNA

A:Residues: 1-538 <SH1>

A:Cross-references: UNIPROT:P03332; UNIPARC:UPI000017424

A:Experimental source: clone pMLV-1

C:Comment: This protein is synthesized as a gag-pol polyprotein.

C:Genetics:

A:Gene: gag

C:Superfamily: mammalian retrovirus gag polyprotein 1

C:Keywords: core protein, inner coat protein, nucleoprotein; polyprotein

F:2-131/Product: core protein p15 #status predicted <C15>

F:132-215/Product: inner coat protein p12 #status predicted <C12>

F:216-478/Product: core shell protein p30 #status predicted <C30>

F:479-534/Product: nucleoprotein p10 #status predicted <C10>

Alignment Scores:

Pred. No.:	1-25e-86	Length:	538
Score:	1376.50	Matches:	283
Percent Similarity:	62.0%	Conservative:	68
Best Local Similarity:	50.0%	Mismatches:	142
Query Match:	48.3%	Indels:	73
DB:	1	Gaps:	13

US-10-723-552-3_COPY_585_2156 (1-1572) x FOMVIM (1-538)

```

QY 1 ATGGGACAGCGGTGACGACCCCTCTTAGTTTGAAGTCTCGACCAATTCGACTGAGTTAAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20

QY 61 TCCAGGCTCATATTTGTCTAGTTTCAAGTAAAGGACCTTGGCAGACTTTCTGTGTC 120
Db 21 ArgileAlaHisAsnGlnSerValAspValLysLysArgTrpValThrPheCysSer 40

QY 121 TCTGAATGGCGGCATTCGATTTGGATGGCCATCAGAGGGACCTTTAATTCGAGATT 180
Db 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60

QY 181 ATCTGGCTGTAAAGCAGTTATTTTTCAGACTGACGACCGCTCTCATCCGATCAGGAG 240
Db 61 IleThrGlnValLysLysValPheSerProGlyProHisGlyHisProAspGlnVal 80

QY 241 CCTATATCTTACGTGGCAAGATTTGGCAGAGGATCTCCGCAATGGGTAAACCATGG 300
Db 81 ProTrpIleValThrTrpGluAlaLeuAlaPheAspProProTrpValLysProPhe 100

QY 301 CTGATTAAGCAAGAACCCAGGTCCCGAATTCCTGGCTCTTGGAGAGAGAAAAACAAC 360
Db 101 ValHis---ProLysProProProProLeuProProSerAlaProSerLeuProLeuGlu 119

QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCGAGATTGAGGAGCCA---CCG 417
Db 120 ProProArgSerThrProProArgSerSerLeuTrpProAlaLeuThrProSerLeuGly 139

QY 418 GCTTGGCCGGAACCCCAA-----GGAGCTCCGGCGGTGAGGACCT-----TCT 438
Db 140 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuIleAspLeuLeuThr 159

QY 439 GTTCCCCCAACCCCTTATCTGGCACAGGTTGCCGCGAGGGACCTTTGGCCCTCCTCT 495
Db 160 GluAspProProProTyrr-----ArgAspProArgProProProSer 173

QY 496 -----GGAGCTCCGGCGGTGAGGACCT-----519
Db 174 AspArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyGluAlaProAspProSerPro 193

QY 520 ---GCTGCAGGAGCTCGAGCGGAGGGCGGCCACCCCGGAGCGGACAGACGAGATCGCG 576
Db 194 MetAlaSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrThrSerGln 213

QY 577 ACATTACCGTCGCGACGTACCGCCCTCCACACCCGGGGGGCCAAATTGACGCGCCCTCCAG 636
Db 577 MetAlaSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrThrSerGln 213

```

RESULT 9

FOMVGV

gag polyprotein - AKV murine leukemia virus

```

Db 214 AlaPheProLeuArgAlaGlyGly-----AsnGlyGln-----LeuGln 226
QY 637 TATTGGCCCTTTCTTCTGAGAGTCTATAATTTGGAAAACCTAACCATCCCTTCTTCG 696
Db 227 TyrTrpProPheSerSerAspLeuTrpAsnTrpLysAsnAsnAsnProSerPheSer 246
QY 697 GAGGATCCCAACGCCCTCACGGGTTGGTGGAGTCCCTTATGTTCTCTCACGAGCTACT 756
Db 247 GluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuIleThrHisGlnProThr 266
QY 757 TGGGATCATTTCTCAACAGCTGCTCCACAGACTCTTCAACCCGAGGAGCGAGAGAATT 816
Db 267 TrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgVal 286
QY 817 CTATTAGAGCTAGAAAAAATGTTCTTGGGCGGACGCGGCGGACCCAGCGGTTCARAA 876
Db 287 LeuLeuGluAlaArgLysAlaValArgGlyAspAspGlyArgProThrGlnLeuProAsn 306
QY 877 GAGATTGACATGGGATTTCCTTAACCTCGCCCGGTTGGGACTACAAACACCGCTGAAGGT 936
Db 307 GluValAspAlaAlaPheProLeuGluArgProAspTrpAspTrpThrThrGlnAlaGly 326
QY 937 AGGAGAGCTTGAATAATCTATCCGACGCTCTGTGGCGGCTCTCCGGGGCGCTCAAGA 996
Db 327 ArgAsnHisLeuValHisTyrArgGlnLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 346
QY 997 CGGCCCACTAATTTGGCTAAGTAAGTAAGTATGATGACGAGGACCGAATGAACCCCTCT 1056
Db 347 SerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer 366
QY 1057 GTTTTCTTGAGAGGCTCTTTGGAAGCCTTCAGGCGGTACACCCCTTTTGTATCCACCTCA 1116
Db 367 AlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp 386
QY 1117 GAGCCCAAAAAGCCTCAGTGGCTTTGGCCCTTTATAGGACAGTCAAGCTTGGATATTAGA 1176
Db 387 ProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly 406
QY 1177 AAGAAGCTTCAGAGACTGGAAGGTTACAGGAGGCTCAGTACGTAGTCTAGTGAAGGAG 1236
Db 407 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 426
QY 1237 GCAGAGAAAGTATTATCAAAAAGGAGACAGAGAAGAAAGGAAACAAAGAAAGAGAGA 1296
Db 427 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgLysArgArg 446
QY 1297 GAAAGAGAGAAAGGAGGAGAAAGA-----CGT 1323
Db 447 GluThrGluGluLysGluGluArgArgThrGluAspGluGlnLysGluLysGluArg 466
QY 1324 AATAAAGCGCAAGAGAAAGATTGACTTAAGATCTTGGCTCAGTGGTTGAGGGAAGAAC 1383
Db 467 AspArgArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys 486
QY 1384 AATACGAAAGAGAGAGAGATTTTAGGAAATTTAGGTACGCGCTCAGACAGTCAGGAA 1443
Db 487 Gln-----AspArgGlnGlyGlyGlu 493
QY 1444 CTGGGCAATAGGACCCCACTCGAAGAGCAATATGTCATATTATTAAAGAAAGAGGACAC 1503
Db 494 ---ArgArgArgSerGlnLeuAspArgAspGlnCysAlaTyrCysLysGluLysGlyHis 512
QY 1504 TGGGCAAGGAGTCCCGGAGAG-----GGAAACAAAGGACCAAG-----1545
Db 513 TrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArgProGlnThrSer 532
QY 1546 ATCTAGCTCTAGAGAA 1563
Db 533 LeuLeuThrLeuAspAsp 538

```

N;Contains: core protein p15; core shell protein p30; inner coat protein p12; nucleoprotein
C;Species: AKV murine leukemia virus
A;Note: host Mus sp. (mouse)
C;Date: 03-Aug-1984 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004
C;Accession: A03933
R;Herr, W.
J. Virol. 49, 471-478, 1984
A;Title: Nucleotide sequence of AKV murine leukemia virus.
A;Reference number: A92995; MUID:84115072; PMID:6319746
A;Accession: A03933
A;Molecule type: genomic RNA
A;Residues: 1-537 <HER>
A;Cross-references: UNIPROT:P03336; UNIPARC:UPI000012B025; GB:J01998; GB:J01999; GB:K000
C;Comment: This protein is synthesized as a gag-pol polyprotein.
C;Genetics:
A;Gene: gag
C;Superfamily: mammalian retrovirus gag polyprotein I
C;Keywords: core protein; inner coat protein; nucleoprotein; polyprotein
F;1-129/Product: core protein p15 #status predicted <C15>
F;130-214/Product: inner coat protein p12 #status predicted <C12>
F;215-477/Product: core shell protein p30 #status predicted <C30>
F;478-537/Product: nucleoprotein p10 #status predicted <C10>

Alignment Scores:
Pred. No.: 2,98e-86 Length: 537
Score: 1371.00 Matches: 282
Percent Similarity: 61.6% Conservative: 72
Best Local Similarity: 49.0% Mismatches: 129
Query Match: 48.1% Indels: 92
DB: 1 Gaps: 14

US-10-723-552-3_COPY_585_2156 (1-1572) x FOMGV (1-537)

Qy 1 ATGGGACAGCGGTGACGACCCCTCTAGTTGACTCTCGACCAATTGGACTGAAGTTAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGluHisTrpGluAspValGln 20

Qy 61 TCCAGGGCTCAATAATTCAGTTCAGTTAAGAGGACCTTGGCAGACTTTCGTGTC 120
Db 21 ArgileAlaSerAsnGlnSerValAspValLysLysArgArGtrpValThrPheCysSer 40

Qy 121 TCTGAATGCCACATTCGATTTGGATGGCCATCAGAGGGGACCTTTAAATTCGAGATT 180
Db 41 AlaGluTrpProThrPheGlyValGlyTrpProGlnAspGlyThrPheAsnLeuaspile 60

Qy 181 ATCTGGCTGTAAAGCAGTTATTTTCAGACTGGAACCCGGCTCTCATCCCGATCAGAG 240
Db 61 IleLeuGlnValLysSerLysValPheSerProGlyProHisGlyHisProAspGlnVal 80

Qy 241 CCCTATATCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATGGTGTAAACCATGG 300
Db 81 ProTyrIleValThrTrpGluAlaIleAlaTyrGluProProProTrpValLysProPhe 100

Qy 301 CTGAATAGCCAAAGACCGGTCCCGAATTCGTGCTCTTGGCTTTGGAGAAAAACAAACAC 360
Db 101 ValSerProLysLeuSerProSerProThrAlaProIleLeu-----Pro 115

Qy 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCAT-----ATCTACCCCGAGATTGAGAGCCA 414
Db 116 SerGlyProSerThrGlnProProProArgSerAlaLeuTyrProAlaLeu---ThrPro 134

Qy 415 CCGGCTTGGCCGAACCCCAA----- 435
Db 135 SerIleLysProArgProSerLysProGlnValLeuSerAspAsnGlyProLeuIle 154

Qy 436 -----TCTGTTCCCCACCCCTTATCTGGCACAGGT----- 468
Db 155 AspLeuLeuSerGluAspProProProTyrGlyGlyGlnGlyLeuSerSerSerAspGly 174

Qy 469 -----CCCGCAGGGGACCCCTTTGGCCCTCTCTGGAGCTCCGGCGGTG 510
Db 175 AspGlyAspArgGluAlaThrSerThrSerGluIleProAlaProSerProIleVal 194

Qy 511 -----GAGGGACCTGCTCAGGAGACTCGGAGCCGAGGCGGCC 549
Db 195 SerArgLeuArgGlyLysArgAspProProAlaAlaAspSerThrThrSerArg----- 212
Qy 550 ACCCGGAGCGGACAGACGATCGACATTACCGTCTGCGACGTACGGCCCTCCACACA 609
Db 213 -----AlaPheProLeuArgLeuGlyGly----- 220
Qy 610 CCGGGGGCCCAATTGTCAGCCCCCTCCAGTATTGGCCCTTTCTCTGCGAGATCTCTAAT 669
Db 221 ---AsnGlyGln-----LeuGlnTyrTrpProPheSerSerAspLeuTyrAsn 236
Qy 670 TGGAAAACTAACCATCCCTTTCTCGAGGATCCCAACGCCCTCAGCGGGTTCGTGAG 729
Db 237 TrpLysAsnAsnAsnProSerPheSerGluAspProGlyLysLeuThrAlaLeuIleGlu 256
Qy 730 TCCCTTATGTTCTCTCCAGCCCTTACCTGGGATGATTGTCAACAGCTGCTCAGACACTC 789
Db 257 SerValLeuThrThrHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlyThrLeu 276
Qy 790 TTCAACACCGGAGCGAGAGAGATTCTATTAGAGCTAGAAAAATGTTCTCTGGGGCC 849
Db 277 LeuThrGlyGluGlyLysGlnArgValLeuLeuGluAlaArgLysAlaValArgGlyAsn 296
Qy 850 GACGGGCGACCCACCGGTTGCCAAATCAGATTGACATGGGATTTCCCTTAACTCGCCCC 909
Db 297 AspGlyArgProThrGlnLeuProAsnGluValAspAlaAlaPheProLeuGluArgPro 316
Qy 910 GGTGGAGCTACACACGGCTGAAGGTAGGAGAGCTTGAATAATCTATCGCAGGCTCTG 969
Db 317 AspTrpAspTyrThrThrGlnArgGlyArgAsnHisLeuValLeuTyrArgGlnLeuLeu 336
Qy 970 GTGGCGGCTCCGGGGCGCTCAAGACGGCCCACTAATTTGGCTTAAGTAAAGAGAGTG 1029
Db 337 LeuAlaGlyLeuGlnAsnAlaGlyArgSerProThrAsnLeuAlaLysValLysGlyIle 356
Qy 1030 ATGACGAGCGACCGAATGAACCCCTCTGTTTCTTCAGAGGCTCTTGGAAAGCTTCAGG 1089
Db 357 ThrGlnGlyProAsnGluSerProSerAlaPheLeuGluArgLeuLysGluAlaTyrArg 376
Qy 1090 CGGTACACCCCTTTGATCCCATCCTCAGAGCCCAAAAGCCTCAGTGGCTTTGGCCCTTT 1149
Db 377 ArgTyrThrProTyrAspProGluAspProGlyGlnGluThrAsnValSerMetSerPhe 396
Qy 1150 ATAGACAGCTCAGCTTGGATTATAGAAAGAGCTTCAGACAGCTGAGAGGCTTACAGGAG 1209
Db 397 IleTrpGlnSerAlaProAspIleGlyArgLysLeuGluArgLeuLysAspLeuLysSer 416
Qy 1210 GCTGAGTTACGTGATCTAGTGAAGGAGCGACAGAAAGTATATTACAAAGGAGACAGAA 1269
Db 417 LysThrLeuGlyAspLeuValArgGluAlaGluArgIlePheAsnLysArgGluThrPro 436
Qy 1270 GAAGAAAGGGAACAAAGAAAGAGAGAGAAAGAGAAAGGAGGAGGAAAGA----- 1320
Db 437 GluGluArgGluGluArgValArgArgGluThrGluGlyLysGluArgArgArgAla 456
Qy 1321 -----CGTAATAAAGCGCAAGAGAGAAATTCGACTAAGATC 1356
Db 457 GluGluGlnLysGluLysGluArgAspArgArgHisArgGluMetSerLysLeu 476
Qy 1357 TTGGCTGCGAGTGGTTGAAGGGAAGCAATACGGAAGAGAGAGAGATTTTAGGAAAAAT 1416
Db 477 LeuAlaThrValSerGlyGln-----ArgArgProGlnLeuAspLysAspGln 487
Qy 1417 AGGTGAGCCCTAGACAGTCTAGGGAACCTGGGCAATAGGACCCCACTCGACAGGACCAA 1476
Db 488 ArgGlnGlyGlyGluArg-----ArgArgProGlnLeuAspLysAspGln 502
Qy 1477 TGTGCTATTCTAAGAAAGAGGACACTGGCAAGGAACCTCCCAAG-----GGA 1530
Db 503 CysAlaTyrCysLysGluLysGlyHisTrpAlaLysAspCysProLysLysProArgGly 522
Qy 1531 AACAAAGGACCAAGG-----ATCCTAGTCTTAGAAGAA 1563

QY	490	CCTCTGTGA-----GCTCCGGCGGTGGAG	513
DB	172	ProAspGlyAspGlyAspArgGluGluAlaThrTyrThrSerGluIleProAlaProSer	191
QY	514	GGACCTGCTGCAGGACTCGGAGCGCGAGGGCGCCACCCCGGAGCGGACAGACGAGATC	573
DB	192	ProMetValSerArgLeuAargGlyLysArgAspProProAlaAlaAspSerThrThrSer	211
QY	574	CGGCACATTTACCGCTCGCAGCATGTCGGCCCTCCACACCGGGGGGCCCAATTGCAGCCCTC	633
DB	212	ArgAlaPheProLeuArgLeuGlyGly-----AsnGlyGln-----Leu	224
QY	634	CAGTATTGGCCCTTTCTTCGACAGACTCTATAATTGGAAAACTAACCATCCCTTC	693
DB	225	GlnTyrTrpProPheSerSerSerAspLeuTyrAsnTrpLysAsnAsnAsnProSerPhe	244
QY	694	TCGAGAGATCCCAACCCCTCAGGGGTGTGGAGTCCCTTATGTCTCTCACCCAGCCT	753
DB	245	SerGluAspProGlyLysLeuThrAlaLeuIleGlnSerValLeuThrThrHisGlnPro	264
QY	754	ACTTGGGATGATTTCAACAGCTGCTCGACACACTCTTCAACAACGAGGAGCGAGAGAGA	813
DB	265	ThrTrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArg	284
QY	814	ATTCTATTAGAGCTAGAAAAATGTTCTCGGGCCGACGGGCGACCCACGCGGTGC	873
DB	285	ValLeuLeuGluAlaArgLysAlaValArgGlyAsnAspGlyArgProThrGlnLeuPro	304
QY	874	AATCAGATTGACATGGGATTTCCCTTAACTCGCCCCGGTTGGGACACTACAACGGCTG	933
DB	305	AsnGluValAsnSerAlaPheProLeuGluArgProAspTrpAspTyrThrTrpGlu	324
QY	934	GGTAGGAGAGCTTGAATAATCTATCGCAGGCTCTGTGTGGCGGTCTCCGGGCGCTCA	993
DB	325	GlyArgAsnHisLeuValLeuTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGly	344
QY	994	AGACGGCCCACTAATTTGGCTAAGGTAAAGAAAGTATGTCAGGACCGAATGAACCCCC	1053
DB	345	ArgSerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerPro	364
QY	1054	TCGTCTTTCTTCAGAGGCTCTTGGAGCCCTTCAGGGCGGTACACCCCTTTTGATCCC	1113
DB	365	SerAlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGlu	384
QY	1114	TCAGAGCCCAAAAGCCCTCAGTGGCTTTGGCCCTTTATAGCAGCAGCTTGGATATT	1173
DB	385	AspHisGlyGlnThrSerValSerMetSerPheIleTrpGlnSerAlaProAspIle	404
QY	1174	AGAAAGACTTCAGACACTGGAAGGTTACAGGAGCTGAGTTACGTGATCTAGTGAAG	1233
DB	405	GlyArgLysLeuGluArgLeuGluAspLeuLysSerLysThrLeuArgAspLeuValArg	424
QY	1234	GGCGCAGAAAGTATATTAACAAAGGAGACAGAGAAGAAAGGCAACAAGAAAGAG	1293
DB	425	GluAlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgPheArg	444
QY	1294	AGAAAGAGAGAAAGGAGGAGAAAG-----AArgInAspArgGlnGlyGluArg-----	1320
DB	445	ArgGluThrGluGlnAsnGluLysArgArgAlaGluAspGluGlnArgGluLysGlu	464
QY	1321	CGTAATAACCGCAAGAGAAATTTGCATAGATCTTCGCTGCAGTGGTTGAAGGAAA	1380
DB	465	ArgAspArgArgGlnArgGluMetSerLysLeuLeuAlaThrValValThrGlyGln	484
QY	1381	AGCAATACGGAAGACAGAGAGATTTTAGAAATTTAGTTCAGGCCCTAGACAGTCAGGG	1440
DB	485	-----AArgInAspArgGlnGlyGluArg-----493	
QY	1441	AACCTGGCAATAGGACCCACTCGACAGGACCAATGTGCATATTTGTAAGAAAGAGCA	1500
DB	494	-----LysArgProGlnLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly	510
QY	1501	CACCTGGCAAGGAACCTGCCCAAGAG-----GGAAACAAGGACCAAGG-----	1545


```
QY 1381 AGCAATACGGAAGAGAGAGAGATTTTAGGAAATTTAGTCTAGGCGCCCTAGACAGTCAGGG 1440
Db 485 -----ArgGlnAspArgGlnGlyGlyGluArg----- 493
QY 1441 AACCTGGCAATAGGACCCCACTGCACAGGACCAATGTGTCATATTTGTAAGAAAGAGGA 1500
Db 494 -----ArgArgProGlnLeuAspGlnCysAlaTyrCysLysGluLysGly 510
QY 1501 CACTGGGCAAGAACTGCCCAAGAAG-----GGAACCAAGACCAAG----- 1545
Db 511 HisTrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArgProGlnThr 530
QY 1546 ---ATCTAGCTCTAGAGAA 1563
Db 531 SerLeuLeuThrLeuAspAsp 537

RESULT 13
FOMVME
gag polyprotein - murine leukemia virus (strain DEF27)
N:Contains: core protein p15; core shell protein p30; inner coat protein p12; nucleoprotein
C:Species: murine leukemia virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: B40416
R:Chattopadhyay, S.K.; Sengupta, D.N.; Fredrickson, T.N.; Morse III, H.C.; Hartley, J.W.
J. Virol. 65, 4232-4241, 1991
A:Title: Characteristics and contributions of defective, ecotropic, and mink cell focus-
A:Reference number: A40416; MUID:91303677; PMID:1649328
A:Accession: B40416
A:Molecule type: DNA
A:Residues: 1-536 <CHA>
A:Cross-references: UNIPROT:P29168; UNIPARC:UPI000012B028; GB:M64096
C:Genetics:
A:Gene: gag
C:Superfamily: mammalian retrovirus gag polyprotein I
C:Keywords: core protein; inner coat protein; nucleoprotein; polyprotein
F:1-129/Product: core protein p15 #status predicted <PI>
F:130-213/Product: inner coat protein p12 #status predicted <ICP>
F:214-476/Product: core shell protein p30 #status predicted <CSP>
F:477-536/Product: nucleoprotein p10 #status predicted <NPI>

Alignment Scores:
Pred. No.: 1,216-84 Length: 536
Score: 1347.50 Matches: 282
Percent Similarity: 62.3% Conservative: 69
Best Local Similarity: 50.1% Mismatches: 127
Query Match: 47.3% Indels: 85
DB: 1 Gaps: 15

US-10-723-552-3_COPY_585_2156 (1-1572) x FOMVME (1-536)
QY 1 ATGGACAGACGGTGACGACCCCTCTTAGTTTGACTCTCGACCAFTTGACTGAAATTAA 60
Db 1 MetGlyGlnThrIleThrProLeuSerLeuThrLeuGluHisTrpArgAspValGln 20
QY 61 TCCAGGGCTCATATTTGCTAGTTTCAAGGGACCTTGGCAGACTTTCGTGTC 120
Db 21 CysIleAlaSerAsnGlnSerValAspValLysArgArgTrpValThrPheCysSer 40
QY 121 TCTGAATGCGACATTCGATCTCGATGCGCATCAGAGGGGACCTTTAAATTCAGATT 180
Db 41 ValGluTrpProSerPheAspValGlyTrpProLeuAspGlyThrPheAsnLeuAspIle 60
QY 181 ATCTCTGGTGTAAAGCAGTATTTTTCAGACTGGACCGGGCTCTCATCCGATCAGGAG 240
Db 61 IleLeuGlnValLysSerLysValPheCysProGlyProHisGlyHisProAspGlnVal 80
QY 241 CCTATATCTTACGTGGCAAGATTTGGCAGAGGATCTCCGCCATGGTGAACCATGG 300
Db 81 ProTyrIleValThrTrpGluAlaLeuAlaTyrHisProProTrpValLysProPhe 100
QY 301 CTGAAT---AGCCA-----AGAAAGCCAGGTCCC----- 327
```

```
Db 101 ValSerProLysProPheProLeuSerThrLeuProPheSerProProGlyProSerAla 120
QY 328 -----CGAATTCGTGCTCTTGAGAGAGAAACAAACAC 360
Db 121 HisProProSerArgSerAspLeuTyrThrAlaLeuLeuProSerIleLysThrLysPro 140
QY 361 TCGGCTGAAAAAGTCAAGCCCTCT-----CCTCATATCTACCCCGAGATTTGAGGAG 411
Db 141 ProLysSerArgValLeuProThrAsnGlyGlyProLeuIleAspLeuLeuThrGluAsn 160
QY 412 CCA-----CCGCTTGGCGCGAGAACCCCAATCTGTT 441
Db 161 ProProAsnLeuGlyGluGlnGlyProProLeuProLysGlyProValLysLysArgArg 180
QY 442 CCCCACCCCTTATCTGGCACAGGTGCCGAGGGGACCTTTGCCCTCTCTGGAGCT 501
Db 181 ProProProPro-----ArgTyrSerProProGly--- 190
QY 502 CCGCGCGTGGAGGACCTGCTGCAGGACTCGAGGCGGAGGCGCCACCCCGAGCGG 561
Db 191 ProMetVal-----SerArgLeuArgGlyAsnArgAspProProAla 204
QY 562 ACAGACGAGATCCGACA-----TTACCGCTGGCAGCTACGCGCCCTCCACACCGGG 615
Db 205 AlaAspSerThrThrSerArgAlaPheProLeuArgLeuGlyGly-----Asn 220
QY 616 GGCCAAATGGACCCCTCCAGTATTCGCTTCTTCAGATCTCTATAATTGAAA 675
Db 221 GlyGln-----LeuGlnTyrTrpProPheSerSerSerAspLeuTyrAsnTrpLys 237
QY 676 ACTAACCATCCCTTTCTCGAGGATCCCAACCGCTCAGCGGTGTGGTGGATCCCT 735
Db 238 AsnAsnAsnProSerPheSerGluAspProGlyLysLeuThrAlaLeuIleGluSerVal 257
QY 736 ATGTTCTCTCACCACTTCTGGGATGTGTCAACAGCTGTGCGACACTCTTCACA 795
Db 258 LeuThrThrHisGlnProThrTrpAspCysGlnGlnLeuLeuGlyThrLeuLeuThr 277
QY 796 ACCGAGGACGAGAGAAATCTATTAGAGCTAGAAAATGTTCTCGGCGCGAGGG 855
Db 278 GlyGluGluLysGlnArgValLeuLeuGluAlaArgLysAlaValArgLysAspGly 297
QY 856 CGACCCACGCGGTGCAAAATGAGATTGACATGGGATTTCCCTTAACTCGCCCCGTTGG 915
Db 298 ArgProThrGlnLeuProAsnGluValAsnSerAlaPheProLeuGluArgProAspTrp 317
QY 916 GACTACAAACCGCTGAAGGTAGGAGAGCTTGAATAATCTATCCGAGGCTCTGTTGGCG 975
Db 318 AspTyrThrThrProGluGlyArgAsnHisLeuValLeuTyrArgGlnLeuLeuAla 337
QY 976 GGTCTCCGGGCGCTCAAGACGCGCCCACTAATTTGGCTAGGTAAGAGAGTGTATGAG 1035
Db 338 GlyLeuGlnAsnAlaGlyArgSerProThrAsnLeuAlaLysValLysGlyIleThrGln 357
QY 1036 GGACCGAATGAACCCCTCTGTTTCTTGAGAGGCTCTTGGAGGCTCTTGGCGGCTAC 1095
Db 358 GlyProSerGluSerProSerAlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyr 377
QY 1096 ACCCTTTTGTATCCCACTCAGAGCGCCCAAAACGCTCAGTGGCTTTGGCCCTTATAGGA 1155
Db 378 ThrProTyrAspProGluAspProGlyGlnGluThrAsnValSerMetSerPheIleTrp 397
QY 1156 CAGTCAGCTTGGATATTAGAAAGACTTCAGAGACTGGAAGGTTTACAGAGGCTGAG 1215
Db 398 GlnSerAlaProAspIleGlyArgLysLeuGluArgLeuGluAspLeuLysSerIysThr 417
QY 1216 TTACTGTATCTAGTGAAGGACGAGAGAAATATATTACAAAAGGGGACAGAGAGAA 1275
Db 418 LeuGlyAspLeuValArgGluAlaGluLysIlePheAsnLysArgGluThrProGluGlu 437
QY 1276 AGGGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 438 ArgGluGluArgIleArgArgGluThrGluGluLysGluGluArgArgArgAlaGluAsp 457
```

81	ProTyrIleValThrTrpGluAlaLeuValTyrHisProProProTrpValLysProPhe	100
301	CTGAAT---AAGCCA-----AGAAAGCCAGGTCCCGCAATT	333
101	ValSerProLysProPheProLeuSerThrLeuProPheSerProGlyProSerAla	120
334	CTGGCTTTGGAGAAAAACAACACTCGGT-----GAAAAGTCAAGCCCTCTCTCT	387
121	HisProProSerArgSerAspLeuTyrThrAlaLeuIleProSerLysProProLysSer	140
388	CATATCTACCCCGAGATTGAGGACCAACCGCTTGGCGGAAACCCCAATCTGTTCGCCCA	447
141	ArgValLeuProProAsnGlyGlyProLeuIleAspLeuThrGluAsnLeuProAsn	160
448	CCCCCTTATCTGGCACAGGTGCCGCGAGGGACCTTTGCCCTCTCTGAGCTCCGCGC	507
161	LeuProProLeuSerLysGlyProValLysLysArgArgProProProArgTyrSer	180
508	GTGAGGAGACT--GCTGCAGGACTCGGAGCGCGAGGGCGCCACCCCGGAGCGGACA	564
181	ProProAsnProMetGluSerArgValArgLysArgArgAspProProAlaAlaAspSer	200
565	GACGAGATCGACATTACCGCTCGGCACGTACGGCCCTCCACACGGGGGGCCCAATTG	624
201	ThrSerSerGlnAlaPheProLeuArgMet-----GlyGlyAspGly	214
625	CAGCCCCCTCAGATTGGCCCTTTCTTCGCAGACTCTATAATTGGAAACTAAACCAT	684
215	Gln---LeuGlnTyrTrpProPheSerSerSerAspLeuTyrAsnTrpLysAsnAsn	233
685	CCCCCTTTCTGGAGATCCCCAACCCCTCACGGGGTGTGGAGTCCCTTATGTTCTCT	744
234	ProSerPheSerGluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuThrThr	253
745	CACCAGCTACTTGGATGTACTCAACAGCTCTCGACAGACATCTTCACACCGAGGAG	804
254	HisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGlu	273
805	CGAGAGAGAATCTATTAGAGGCTAGAAAAAATGTTCTTGGGGCCGACGGGCCACCCACG	864
274	LysGlnArgValLeuLeuGluAlaArgLysAlaValArgLysAsnAspGlyArgProThr	293
865	CGGTTCGAAATGAGATTGACATGGGATTTCCTTAACTCGCCCGGTGGGACTACAAC	924
294	GlnLeuProAsnGluValAsnSerAlaPheProLeuGluArgProAspTrpAsnTyrSer	313
925	ACGGCTCAGGTAGGAGGACCTTGAAATCTATCGCAGCTCTGTGGCGGTCTCTCCG	984
314	ThrProGluGlyArgAsnHisLeuValLeuTyrArgGlnLeuLeuLeuAlaGlyLeuHis	333
985	GGCGCTCTAAGACGGCCCACTAATTTGGCTTAAGTAAGAGAGTGTACGAGGACCGAAT	1044
334	AsnAlaGlyArgSerProThrAsnLeuAlaLysValLysArgLysLeuGlnGlyProAsn	353
1045	GAACCCCTCTCTTTTCTTGAGAGCTCTTGGAAGCCCTTACGGGGTACACCCCTTTT	1104
354	GluSerProSerAlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyr	373
1105	GATCCCACTCAGAGGCCCAAAAGCCTCAGTGGCTTTGGCTTTTATAGCAGTCAAGCC	1164
374	AspProGluAspProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAla	393
1165	TTGCATTATAGAAGAAGCTTCACAGACTGGAAGGTTACAGGAGGCTGAGTTACGTGAT	1224
394	ProAspIleGlyArgLysLeuGluArgLeuGluAspLeuLysSerLysThrLeuGlyAsp	413
1225	CTAGTGAAGGAGGACAGAAAGTATTATCAAAAGGGGACAGAGAAGAAAGGGAAACA	1284
414	LeuValArgGluAlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGlu	433
1285	AGAAAAGAGAGAGAGAGAGAAAGGAGGAAAGA-----	1320
434	ArgIleArgArgGluThrGluGluLysGluGluArgArgAlaGluAspGluGlnArg	453

```
QY 1321 -----CCTAATAACCGCAGAGAGAAATTTGACTAAGATCTTGGCTGCAGTGGTT 1371
Db 454 GluLysGluArgAspArgArgHisArgGluMetSerLysPheLeuAlaThrValVal 473
QY 1372 GAAGGGAAGCAATACGGAAGAGAGAGAGATTTTACGAAATTTAGGTCCAGGCCCTAGA 1431
Db 474 ThrGlyGln-----ArgGlnAspArgGlnGlyGlu 484
QY 1432 CAGTCAGGGAACCTGGGCAATAGGACCCACTCGAAGAGCAATGTGCAATTTGTATAA 1491
Db 485 Arg-----ArgArgProGlnLeuAspGluAspGlnCysAlaTy-CysLys 499
QY 1492 GAAGAGGACACTGGGCAAGNACTGCCCCAGAG-----GGNAACAAAGGACCAGG 1545
Db 500 GluLysGlyHisTrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArg 519
RESULT 15
FOMVVB
gag polyprotein - baboon endogenous virus
N:Contains: core protein p15; core shell protein p30; inner coat protein p12; nucleoprotein
C:Species: baboon endogenous virus strain M7
A:Note: host Papio sp. (baboon)
C:Date: 19-Feb-1984 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004
C:Accession: A03939
R:Itamura, T.
J. Virol. 47, 137-145, 1983
A:Title: Provirus of M7 baboon endogenous virus: nucleotide sequence of the gag-pol region
A:Reference number: A03939; MUID:83241915; PMID:6408267
A:Accession: A03939
A:Molecule type: DNA
A:Residues: 1-532 <TM>
A:Cross-references: UNIPROT:P03341; UNIPARC:UPI0000174A29
C:Comment: This protein is synthesized as a gag-pol polyprotein.
C:Genetics:
A:Gene: gag
A:Superfamily: mammalian retrovirus gag polyprotein 1
C:Keywords: core protein; inner coat protein; nucleoprotein; polyprotein
F:1-111/Product: inner coat protein p12 #status predicted <C12>
F:112-218/Product: core protein p15 #status predicted <C15>
F:219-472/Product: core shell protein p30 #status predicted <C30>
F:473-532/Product: nucleoprotein p10 #status predicted <C10>
Alignment Scores:
Pred. No.: 3,29e-82 Length: 532
Score: 1312.00 Matches: 279
Percent Similarity: 64.0% Conservative: 78
Best Local Similarity: 50.0% Mismatches: 141
Query Match: 46.1% Indels: 60
Db: 1 Gaps: 15
US-10-723-552-3_COPY_585_2156 (1-1572) x FOMVVB (1-532)
QY 1 ATGGGACAGAGCGTGACGACGCCCTCTTAGTTTGACTCTCGACCATTTGGACTGAAGTTAA 60
Db 1 MetGlyGlnThrLeuThrProLeuSerLeuThrLeuThrHisPheSerAspValArg 20
QY 61 TCCAGGCTCATAATTTGTCAGTTTTCAGTTTGAAGAGGACCTTGGCAGACTTTCTGTCTC 120
Db 21 AlaArgAlaHisAsnLeuSerValGlyValArgLysGlyArgTrpGlnThrPheCysSer 40
QY 121 TCTGAA-----TGCCCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTCGAG 177
Db 41 SerGluValHisProSerCysArgValAla-----ArgAspGlyThrPheAspLeuSer 58
QY 178 ATTATCTCGGCTGTTAAAGACGATTATTTTCAGACTGACCGCGGCTCTCATCCCGATCAG 237
Db 59 ValIleLeuGlnValLysThrLysAspMetAspProGlyProHisGlyHisProValPro 78
QY 238 GAGCCCTATATCTTACGTGGCAAGATTTGGCAGAGGATCTCCGCCATGGGTTAAACCA 297
Db 79 SerGlyTyrlleIleThrTrpValAspLeuAspGlyAsnProProProTrpGlyLysPro 98
```

```
QY 298 TGGCTGAATAAGCAAGAACCCAGGTCGCCGAATTTCTGGCTCTTGGAGAGAAAAACAAA 357
Db 99 PheLeuHisThrProSerThrSerLysSerThrLeuLeuAlaLeuGluValProLysAsn 118
QY 358 CACTCGGCTGAAGAAGTCAAGCCCTCTCTCATATCTACCCCGAG----- 402
Db 119 ArgThrLeuAspProProLysPro-----ValLeuProAspGluSerGlnGlnAsp 135
QY 403 -----ATTGAGGAGCCA---CCGCTTGGCCGGAACCCCAATCTGTTCGCCACCCCT 453
Db 136 LeuLeuPheGlnAspProLeuProHisProProHisAsnProLeuLeuGluProProPro 155
QY 454 TATCTGGCACAGGCTGCCGCGAGGGACCCCTTTCCTCTCTGGAGCTCCGGCGGTG--- 510
Db 156 TyrAsnSerProSerProProThrProSerAlaProThrProSerSerLeuValSer 175
QY 511 -----GAGGACCTGCTCGAGGAGACTCGGACCGGAGGGCGCCACCCCG 555
Db 176 SerSerThrProProSerSerProAlaProProGluLeuThrProArg-----ThrPro 193
QY 556 GAGCGGACAGCAGAGATC----- 573
Db 194 ProGlnThrProArgLeuArgLeuArgAlaGluGlyGlnAspGlyProPheHisLeu 213
QY 574 -----GCGACATTACCGCTCGCACAGCTACGGCCCTCCACACCGGGGGGCCAATTGCAG 627
Db 214 GlnSerSerLeuPheProLeuArgThrValAsnArgThrValLys-----Arg 229
QY 628 CCCCTCCAGTATTCGCCCTTTTCTTCTGCAGACTCTATATAATCGAAACCTAACCATCCC 687
Db 230 ThrIleGlnTyTrpProPheSerAlaSerAspLeuTyAsnTrpLysThrHisAsnPro 249
QY 688 CCTTCTCGGAGGATCCCAACCGCTCACGGGGTGTGGAGTCCCTATATTCTCTCTCAC 747
Db 250 SerPheSerGlnAspProGlnAlaLeuThrSerLeuIleGluSerIleLeuLeuThrHis 269
QY 748 CAGCTACTTGGGATGATTGTCAACAGCTGTCGAGACACTCTTTCACAAACCGAGAGCGA 807
Db 270 GlnProThrTrpAspAspCysGlnGlnLeuLeuGlnValLeuLeuThrThrGluGluArg 289
QY 808 GAGAGAATTCATTAGAGGCTAGAAAAAATTTCTCGGGCGCGAGCGGCGCGCCACCGCGG 867
Db 290 GlnArgValLeuLeuLeuGluAlaArgLysAsnValProGlyProGlyGlyLeuProThrGln 309
QY 868 TTGCAAAATGAGATTGACATGGGATTTCCCTTAATCGCCCGGTTGGAGTACACACAG 927
Db 310 LeuProAsnGluIleAspGluGlyPheProLeuThrArgProAspTrpAspTrpGluThr 329
QY 928 GCTGAAGGTAGGAGAGCTTGAATAATCTATCCAGAGGCTCTGGTGGCGGGTCTCCGGGCG 987
Db 330 AlaProGlyArgGluSerLeuArgIleTyArgGlnAlaLeuLeuAlaGlyLeuLysGly 349
QY 988 GCCTCAAGACGGCCCACTAATTTGGCTAAGGTAAGAGAAAGTGATCGAGGAGCCGAATGAA 1047
Db 350 AlaGlyLysArgProThrAsnLeuAlaLysValArgThrIleThrGlnGlyLysAspGlu 369
QY 1048 CCCCTCTCTGTTTCTTTCAGAGGCTCTTGAAGCTTTCAGCGGTCACCCCTTTTCAT 1107
Db 370 SerProAlaAlaPheMetGluArgLeuLeuGluGlyPheArgMetTyThrProPheAsp 389
QY 1108 CCCACCTCAGAGGGCCCAAAAAGCCTCAGTGGCTTTGGCTTTATAGGACAGTCAGCCCTTG 1167
Db 390 ProGluAlaProGluHisLysAlaThrValAlaMetSerPheIleAspGlnAlaLeu 409
QY 1168 GATATTAGAAAGAGCTTTCAGAGACTGGAAGGGTTTACAGGAGGCTTGACTTACGTGATCTA 1227
Db 410 AspIleLysGlyLysLeuGlnArgLeuAspGlyIleGlnThrHisGlyLeuGlnGluLeu 429
QY 1228 GTGAAGGCGCAGAGAAAGTATATTAACAAAGGAGCAGAGAGAGAAAGGACAAAGA 1287
Db 430 ValArgGluAlaGluLysValTyAsnLysArgGluThrProGluGluArgGluAlaArg 449
QY 1288 AAAGAGAGAGAAAGAGAGAGAAAGGAGAAAGAGCTTAATAAAGCGCAAGAGAAATTTG 1347
```


This Page Blank (uspto)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 12:52:14 ; Search time 65.4051 Seconds
(without alignments)
3391.447 Million cell updates/sec

Title: US-10-723-552-3_COPY_585_2156

Perfect score: 2848

Sequence: 1 ATGGGACAGCGTGACGAC.....CTCTAGAGAGATAAAGAT 1572

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abas/ABSSWEB/spool/US10723552/runat 14022006 125141 12876/app query.fasta_1
-DB-Uniprot -QFWT-fastan -SUFFIX=rup -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN 1.1 808 @runat 14022006 125141 12876 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -SPBLOC=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2817	98.9	524	2	O62706 sus scrofa
2	2813	98.8	524	2	O62704 sus scrofa
3	2812	98.7	524	2	O5QG09 9GAMR
4	2774	97.4	524	2	O4VF23 9GAMR
5	2757	96.8	524	2	O6UJ23 9GAMR
6	2747	96.5	524	2	O6T3V4 9GAMR
7	2747	96.5	524	2	O8UMA0 9GAMR
8	2745	96.4	524	2	O90RM0 9GAMR
9	2725	95.7	524	2	O6UJ25 9GAMR
10	2724	95.6	524	2	O6UJ22 9GAMR
11	2705	95.0	664	2	O9N177 9GAMR
12	2705	95.0	1718	2	O8J4V8 9GAMR
13	2705	95.0	1720	2	O8J4V6 9GAMR
14	2705	95.0	2378	2	O9Q1X4 9GAMR
15	2702	94.9	524	2	O62708 sus scrofa
16	2701	94.8	525	2	O6T3V5 9GAMR

17	2699	94.8	524	2	Q8UMP6 9GAMR
18	2699	94.8	524	2	Q90RL7 9GAMR
19	2699	94.8	2376	2	Q9Q1X3 porcine end
20	2698	94.7	524	2	Q8Q6U8 9GAMR
21	2695	94.6	2376	2	Q9Q1X5 9GAMR
22	2693	94.6	524	2	Q7ZBT7 9GAMR
23	2686	94.3	524	2	Q8UM97 9GAMR
24	2676	94.0	524	2	Q73504 9GAMR
25	2419.5	85.0	525	2	Q8Q6U5 porcine end
26	1733	60.8	522	2	Q9WLL1 mus dunni e
27	1733	60.8	622	2	Q89814 mus dunni e
28	1718.5	60.3	521	2	Q9TTC2 PHACI
29	1672.5	58.7	519	1	GAG_GALV
30	1656.5	58.2	522	2	O70651 GALV
31	1632.5	57.3	511	1	GAG SMSAV
32	1472.5	51.7	536	1	GAG_BAEVM
33	1412.5	49.6	536	2	Q7ZJT8 9GAMR
34	1412.5	49.6	1736	2	Q7ZJT6 9GAMR
35	1411.5	49.6	535	1	GAG_MLVCB
36	1409.5	49.5	536	2	O6YTY0 MOUSE
37	1398.5	49.1	534	2	O9YKA0 9GAMR
38	1396.5	49.0	534	2	O6ORS7 9GAMR
39	1393.5	48.9	1733	2	Q9E7M1 9GAMR
40	1388	48.7	538	1	GAG_MLVF5
41	1387.5	48.7	538	2	O67457 MLVPR
42	1385.5	48.6	538	2	O12373 9GAMR
43	1384.5	48.6	521	2	O7ZL05 9GAMR
44	1384.5	48.6	521	2	O7ZL07 9GAMR
45	1384.5	48.6	538	2	O41249_MLVRA

ALIGNMENTS

RESULT 1

O62706_PIG

ID O62706_PIG PRELIMINARY; PRT; 524 AA.

AC O62706;

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Gag protein.

OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;

OC Sus.

OX NCBI_TaxID=9823;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=miniature swine;

RX MEDLINE=98216827; PubMed=9557749;

RA Akiyoshi D.E., Denaro M., Zhu H., Greenstein J.L., Banerjee P.,

RA Fishman J.A.;

RT "Identification of a full-length cDNA for an endogenous retrovirus of

RT miniature swine.";

RL J. Virol. 72:4503-4507(1998).

DR HSP; P03332; IAGB.

DR SMR; O62706; 7-98.

DR GO; GO:0019028; C:viral capsid; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR GO; GO:0019068; P:viral assembly; IEA.

DR InterPro; IPR003036; Gag_p30.

DR InterPro; IPR000840; G_retro.matrix.

DR InterPro; IPR001878; Znf_CCHC.

DR Pfam; PF01140; Gag_MA; 1.

DR Pfam; PF02093; Gag_p30; 1.

DR Pfam; PF00098; zf_CCHC; 1.

DR SMART; SM00343; Znf_C2HC; 1.

DR PROSITE; PS00158; ZF_CCHC; 1.

SQ SEQUENCE 524 AA; 59548 MW; C2D67640A70898B8 CRC64;

Alignment Scores:

Pred. No.:	1.02e-168	Length:	524
Score:	2817.00	Matches:	524
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	98.9%	Indels:	0
DB:	2	Gaps:	0

US-10-723-552-3_COPY_585_2156 (11-1572) x 062706_PIG (1-524)

Qy	1	ATGGGACAGCGGTGACACACCCCTTTAGTTGACTCTCGACCATTGAGCTGAAGTTAAA	60
Db	1	MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluVallys	20
Qy	61	TCCAGGGCTCATAAATTCGTAGTTTACAGTTAAGAGGACCTTGGCAGACTTTCTGTCTC	120
Db	21	SerArgAlaHisAsnLeuSerValGlnVallysylsGlyProTrpGlnThrPheCysVal	40
Qy	121	TCTGAATGCCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAAATTCGAGATT	180
Db	41	SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle	60
Qy	181	ATCTGGCTGTAAAGAGTTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG	240
Db	61	IleLeuAlaValLysAlaValIlePheGlnThrGlyProGlySerHisProAspGlnGlu	80
Qy	241	CCCTATATCCTTACGTGGCAAGATTGGCAGAGGATCCTCCGCCATGGTGTAAACCATGG	300
Db	81	ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpTrpValysProTrp	100
Qy	301	CTGAATAAGCCAAGAACCCAGGTCCCGCAATTCGGCTCTTGGAGAGAAAAACAACAC	360
Db	101	LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis	120
Qy	361	TCGGCTGAAAAGTCAAGCCCTCTCCTCATATCTACCCCGAGATTGAGGACCCGGCT	420
Db	121	SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProProAla	140
Qy	421	TGGCCGGAACCCCATCTGTTCCCCACCCCTTATCTGGCACAGGGTCCCGCAGGGGA	480
Db	141	TrpProGluProGlnSerValProProProProTyrLeuAlaGlnGlyAlaAlaArgGly	160
Qy	481	CCCTTTGGCCCTCTCTGGAGCTCCGGCGGTGAGGACCTGTCTGAGGACTCGAGCCGG	540
Db	161	ProPheAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg	180
Qy	541	AGGGCGGCCACCCGGAGCGGACAGACGAGATCGGCACATTACCGCTCGGCACGTACGGC	600
Db	181	ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly	200
Qy	601	CCTCCACACCGGGGGCCAAATTGCAGCCCTCCAGTATTGGCCCTTTCTTCTGCAGAT	660
Db	201	ProProThrProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp	220
Qy	661	CTCTATAATGGAAAACATAACATCCCTTTCTCGGAGGATCCCAACGCTCTCACGGGG	720
Db	221	LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly	240
Qy	721	TTGGTGGAGTCCCTTATTTCTCTCACCGACTACTTGGGATGATGTCAACAGCTGCTG	780
Db	241	LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu	260
Qy	781	CAGACACTCTTCACACACCGAGGCGGACAGAGATCTTATTAGCGCTAGAAAATGTT	840
Db	261	GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal	280
Qy	841	CCTGGGCGGACGGCGGACCCACCGGGTTGCAAAATGAGATTGACATGGGATTTCCCTTA	900
Db	281	ProGlyAlaAspGlyArgProThrArgLeuGlnAsnGluIleAspMetGlyPheProLeu	300
Qy	901	ACTGCCCCGGTTGGGATACAAACAGCGCTGAAGGTAGGAGCTTGAAATCTATCGC	960
Db	301	ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg	320

Qy	961	CAGGCTCTGGTGGCGGGTCTCCGGGGCGCTCAAGACGGCCCACTAAATTTGGCTAAGGTA	1020
Db	321	GlnAlaLeuValAlaGlyLeuArgLysAlaSerArgArgProThrAsnLeuAlaLysVal	340
Qy	1021	AGAGAAGTGATGACGGACCCGAATGAACCCCTCTCTGTTTTTCTTGAGAGCTCTTGAA	1080
Db	341	ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuLeuGlu	360
Qy	1081	GCCTTCAGGCGGTACACCCCTTTTGATCCCACTCAGAGGCCCAAAAGCCCTCAGTGCT	1140
Db	361	AlaPheArgArgTyrThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla	380
Qy	1141	TTGGCCCTTTATAGACACAGCTCAGCCCTTGATATTAGAAGAAAGCTTCAGAGACTGGAAGG	1200
Db	381	LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly	400
Qy	1201	TTACAGAGGCTGAGTTACGTGATCTAGTGAAGGAGGACAGAAAGTATATTTACAAAAGG	1260
Db	401	LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg	420
Qy	1261	GAGACAGAGAGAAAGCGGACAAAGAAAGACAGAGAGAAAGAGGAAAGGAGGAAGA	1320
Db	421	GluThrGluGluGluArgGlnArgLysGluArgGluArgGluGluArgGluGluArg	440
Qy	1321	CGTAATAAACCGCAAGAGAGAGAAATTTGACTAAGATCTTGGCTGCAGTGGTTGAAGGAAA	1380
Db	441	ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValGluGlyLys	460
Qy	1381	AGCAATACGGAAGAGAGAGAGATTTTAGGAAAATTTAGGTCAAGCCCTAGACAGTCAGG	1440
Db	461	SerAsnThrGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly	480
Qy	1441	AACCTGGCCATAGACCCCACTCGACAGACCAATCTGCATATTTGTAAAGAAAGAGA	1500
Db	481	AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGly	500
Qy	1501	CAC'TGGGCAAGCAACTGCCCAAGAGGAAACAAAGACCAAGCATCTAGCTCTAGAA	1560
Db	501	HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProArgIleLeuAlaLeuGlu	520
Qy	1561	GAAGATAAAGAT 1572	
Db	521	GluAspLysAsp 524	

RESULT 2

062704_PIG	
ID	062704_PIG PRELIMINARY; PRT; 524 AA.
AC	062704;
DT	01-AUG-1998 (TrEMBLrel. 07, Created)
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Gag protein.
OS	Sus scrofa (Pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC	Sus
OX	NCBI_TaxID=9823;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RX	MEDLINE=98216827; PubMed=9557749;
RA	Akiyoshi D.E., Denaro M., Zhu H., Greenstein J.L., Banerjee P.,
RA	Fishman J.A.;
RT	"Identification of a full-length cDNA for an endogenous retrovirus of
RL	miniature swine.";
RL	J. Virol. 72:4503-4507(1998).
DR	EMBL; AF038599; AAC16761.1; -, mRNA.
DR	HSSP; P03332; lAeB.
DR	SMR; 062704; 7-98.
DR	GO; GO:0019028; C:viral capsid; IEA.
DR	GO; GO:0003676; F:nucleic acid binding; IEA.
DR	GO; GO:0005198; F:structural molecule activity; IEA.
DR	GO; GO:0019068; P:viral assembly; IEA.
DR	InterPro; IPR003036; Gag_p30.

DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_Ma; 1.
DR Pfam; PF02093; Gag_P30; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.
SQ SEQUENCE 524 AA; 59578 MW; 6B54F545098EA1DF CRC64;

Alignment Scores:
Pred. No.: 1,82e-168 Length: 524
Score: 2813.00 Matches: 522
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 98.8% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x 062704_PIG (1-524)

QY 1 ATGGGACAGCGGTGACACCCCTCTTAGTTTGACTCTCGACCAATTGGACTGAAGTTAAA 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTyrThrGluValys 20

QY 61 TCCAGGGCTCATATTCTCAGTTTCAGTTTAAAGAGGACCTTGGCAGACTTCTGTGTC 120
DB 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTyrGlnThrPheCysVal 40

QY 121 TCTGAATGGCGGCACATTTCGATGTGGATGGCCATCAGAGGGGACCTTTAAATTCTGAGATT 180
DB 41 SerGluTyrProThrPheAspValGlyTyrProSerGluGlyThrPheAsnSerGluLeu 60

QY 181 ATCTGGCTGTTAAAGCAGTTATTTTTCAGACTGACCCGCTCTCATCCGATCAGGAG 240
DB 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80

QY 241 CCTATATCTTACGTGTCAGATTTCGACAGGATCCTCGCCATCGGTAAACCATGG 300
DB 81 ProTyrIleLeuThrTyrGlnAspLeuAlaGluAspProProTyrPValLysProTyr 100

QY 301 CTGAATAAGCCAGAAAGCCAGGTCCCGGAATTCCTGGCTCTGGAGAGAGAAAAACAACAC 360
DB 101 LeuAsnLysProAlaGlyLysProGlyProargIleLeuAlaLeuGlyGluLysAsnLysHis 120

QY 361 TCGGCTGAAAAGTCAAGCCCTCTCCTCATATCTACCCGAGATTGAGGAGCCACCGGCT 420
DB 121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluProProAla 140

QY 421 TGGCCGGAACCCCAATCTGTTCCCGCACCCCTTATCTGGCAAGGGTGGCGGAGGGA 480
DB 141 TrpProGluProGlnSerValProProProTyrIleuAlaGlnGlyAlaAlaArgGly 160

QY 481 CCTTTTGCCTCTCTGGAGCTCCCGCGTGGAGGACCTGCTGCAGGGACTCGGAGCGG 540
DB 161 ProPheAlaProProGlyAlaProAlaValGluGlyProSerAlaGlyThrArgSerArg 180

QY 541 AGGGCGCCACCCCGGAGCGGACAGACGAGATTCGCGACATTACCGCTCGCAGCTACGCG 600
DB 181 ArgGlyAlaThrProGluAlaThrAspGluIleAlaThrLeuProLeuAlaThrTyrGly 200

QY 601 CTTCCACACCGGGGGCCAAATTCAGCCCTTCAGTATTCGCGCTTTCTTCTGCGAGAT 660
DB 201 ProProThrProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220

QY 661 CTCTATATATGAAAATCAACATCCCTCTTCGCGAGGATCCCAACGCTCACCGGG 720
DB 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnAlaGluThrGly 240

QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACCAGCTACTTGGGATGATTGTCAACAGCTCTG 780
DB 241 LeuValGluSerLeuMetPheSerHisGlnProThrTyrPheAspCysGlnGlnLeuLeu 260

QY 781 CAGACATCTTCACAACCGAGGCGGAGAGAAATTCATTAGAGGCTAGAAAAATGTT 840

DB 261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
QY 841 CCTGGGCGGACCGGCGGACCCACCGGTTGCAAAATAGATTGACATGGGATTTCCCTTA 900
DB 281 ProGlyAlaAspGlyArgProThrArgLeuGlnAsnGluIleAspMeGlyPheProLeu 300

QY 901 ACTCGCCCGGTTGGGACTCAACACACGCTGAAAGGTAGGAGAGCTTGAAATCTATCGC 960
DB 301 ThrArgProGlyTyrAspTyrAsnThrAlaGluGluArgGluSerLeuLysIleTyrArg 320

QY 961 CAGGCTCTGGTGGGGTCTCCGGGGCCCTCAAGACGCCCACTAAATTTGGCTAAGGTA 1020
DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgProThrAsnLeuAlaLysVal 340

QY 1021 AGAGAAGTGATGACGAGGACCAATGAACCCCTCTGTTTCTTGGAGAGGCTCTTGCAA 1080
DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuLeuGlu 360

QY 1081 GCCTTCAGGCGGTACACCCCTTTTGTATCCCACTCAGAGGCCCAAAAGCCCTCAGTGCT 1140
DB 361 AlaPheArgArgTyrThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380

QY 1141 TTGSCCTTTATAGACAGCTCAGCCCTCGATATTAGAAAGAGCTTCAGAGACTGGAAGG 1200
DB 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400

QY 1201 TTACAGGAGGCTGAGTTACGTGATCTAGTGAAGGAGGAGAGAAAGTATATTACAAAAGG 1260
DB 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420

QY 1261 GAGACAGAAGAAAGGAAACAAAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluArgGluGluArgGluArg 440

QY 1321 GGTAAATAACCGGCAAGAGAAATTTGACTAAGATCTTGGCTGAGTGGTTGAAGGGAAA 1380
DB 441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValGluGlyLys 460

QY 1381 AGCAATACGAAAGAGAGAGAGATTTTAGGAAAATTAGTTCAGGCCCTAGACAGTCAGG 1440
DB 461 SerAsnThrGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480

QY 1441 AACCTGGCAATAGGACCCCACTCGACAGGACCAATGTGCATATTGTAAAGAAAGAGGA 1500
DB 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGly 500

QY 1501 CACTGGCAGGAACTCCCGCAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 501 HisTyrAlaArgAsnCysProLysLysGlyAsnLysGlyProArgIleLeuAlaLeuGlu 520

QY 1561 GAAGATAAAGAT 1572
DB 521 GluAspLysAsp 524

RESULT 3
Q5QG09_9GAMR
ID Q5QG09_9GAMR PRELIMINARY; PRT; 524 AA.
AC Q5QG09_9GAMR
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Gag protein.
OS Porcine endogenous retrovirus C/A.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammaretrovirus; 1-Mammalian type C virus group.
NCBI_TaxID=286072;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A14/220;
RX PubMed=15564496; DOI=10.1128/JVI.78.24.13880-13890.2004;
RA Bartosch B., Stefanidis D., Myers R., Weiss R., Patience C.,
RA Takeuchi Y.;
RT "Evidence and consequence of porcine endogenous retrovirus

```
RT recombination.";
RL J. Virol. 78.13880-13890 (2004).
DR EMBL; AY570960; AAT77166.1; -, Genomic_DNA.
DR SMR; Q5Q0G9; 7-98.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00098; zF_CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS0158; ZF_CCHC; 1.
SQ SEQUENCE 524 AA; 59534 MW; 24805E0EBAA46E1F CRC64;

Alignment Scores:
Pred. No.: 2,1e-168 Length: 524
Score: 2812.00 Matches: 522
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 98.7% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x Q5Q0G9_9GAMR (1-524)
QY 1 ATGGACAGACGGTGACGACCCCTCTTAGTTGACTCTCGACCAATTGACTGAAGTTAAA 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisThrThrGluVallys 20
QY 61 TCCAGGGCTCATAAATTTGCTAGTTAGTTAGAGGACCTTGGCAGACTTCTGTGTC 120
DB 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 40
QY 121 TCTGAATGCCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAAATCTGAGATT 180
DB 41 SerGlnTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
QY 181 ATCTGGCTGTTTAAAGACGATTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
DB 61 IleLeuAlaValLysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
QY 241 CCTATATCTTACGTGGCAGAGATTGGCAGAGGATCTCCGCCATGGGTTAAACCATGG 300
DB 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
QY 301 CTGAATAGCCAAAGAACCCAGGTCCTCCGAAATCTGGCTCTTGGAGAGAAAACAAACAC 360
DB 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCACCGGCT 420
DB 121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProProAla 140
QY 421 TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGACACGGGTCCCGCGAGGGGA 480
DB 141 TrpProGluProGlnSerValProProProProTyrLeuAlaGlnGlyAlaAlaArgGly 160
QY 481 CCTTTGGCTCTCTGGAGCTCCGCGGTGGAGGACCTGTGACAGGACTCGGAGCCGG 540
DB 161 ProPheAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
QY 541 AGGGCGGCCACCCCGAGCGGACAGACGAGATCGGCACATTACCGCTCGCACGTACGGC 600
DB 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly 200
QY 601 CCTCCACACCGGGGGGCAATTGAGCCCTCCAGTATTGGCCCTTCTTCTGTCAGAT 660
DB 201 ProProThrProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220
QY 661 CTCATATAATTGGAAACTAACCATCCCTTTCTCGGAGGATCCCAACAGCCCTCAGGGG 720
```

```
Db 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
QY 721 TTGGTGGAGTCCCTTAATGTTCTCTCACAGGCTACTTGGGATGATTGTCAACAGCTGCTG 780
DB 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
QY 781 CAGACACTCTTCAACACCGAGGACGAGAGAAATCTATTAGAGGCTAGAAAAAATCTT 840
DB 261 GlnThrLeuPheThrThrGluGluArgGluGluIleLeuLeuGluAlaAargLysAsnVal 280
QY 841 CCTGGGCGCCGACGCGGACCCACACGCGTTCGAAATGAGATTGACATGGGATTTCCCTTA 900
DB 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
QY 901 ACTGCCCCGGTGGGACTACAAACACGCGTGAAGGTAGGAGAGCTTGAAATCTATCGC 960
DB 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
QY 961 CAGGCTCTGGTGGCGGCTCTCCGGGGCGCCTCAAGACGGCCCTAAATTTGCGCTAAGGTA 1020
DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgProThrAsnLeuAlaLysVal 340
QY 1021 AGAGAAGTGTATGACAGGACCCGAATGAACCCCTCTCTGTTTTTCTTGAGAGGCTCTTGGAA 1080
DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuLeuGlu 360
QY 1081 GCCTTCAGGCGGTACACCCCTTTGATCCACCTCAGAGGCCCAAAAAGCTCAGTGCT 1140
DB 361 AlaPheArgArgTyrThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
QY 1141 TTGSCCTTTATAGACAGCTCAGCCTTGATATTAAGAAAGAGCTTCAGAGACTGGAAGGG 1200
DB 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLeuGlnArgLeuGluGly 400
QY 1201 TTACAGGAGGCTGAGTTACGTGATCTAGTGAAGAGGCAGAGAAAGTATATTACAAAAGG 1260
DB 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420
QY 1261 GAGACAGAGAAAGAAAGGAAACAAAGAAAGAGAGAGAAAGAGAGAAAGAGGAGGAAGA 1320
DB 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluArgGluGluArgGluGluArg 440
QY 1321 CGTAATAACCGCAAGAGAGAAATTTGACTAAGATCTTGGCTGCAGTGGTTGAAGGGAAA 1380
DB 441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValGluGlyLys 460
QY 1381 AGCAATACGGAAGAGAGAGAGATTTTAGGAAAAATTAGGTCAAGCCCTAGACAGTCAGGG 1440
DB 461 SerAsnThrGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
QY 1441 AACCTGGGCAATAGAGACCCCTCGACAGGACCAATGTGCATATTGTAAAGAAAGAGA 1500
DB 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGly 500
QY 1501 CACTGGGCAAGAACTGCCCCCAAGAGGAAACAAAGGACCAAGGATCCTAGCTCTAGAA 1560
DB 501 HisTrpAlaArgAsnCysProLysGlyAsnLysGlyProArgIleLeuAlaLeuGlu 520
QY 1561 GAAGATAAAGAT 1572
DB 521 GluAspLysAsp 524

RESULT 4
Q4VFZ3_9GAMR PRELIMINARY; PRT; 524 AA.
ID Q4VFZ3_9GAMR PRELIMINARY; PRT; 524 AA.
AC Q4VFZ3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Gag protein.
GN Name=gag;
OS Porcine endogenous retrovirus C/A.
```

OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 CC Gammatretrovirus; 1-Mammalian type C virus group.

OX NCBI_TaxID=286072;
 RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C/A 4;

RA Denner J., Karlas A., Vottele J.;

RT "Nucleotide sequence and extended characterization of a high passage

RT human cell-adapted recombinant PERV-C/A.";

RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY953542; AAY28927.1; -; Genomic_DNA.

DR InterPro; IPR000840; G_retro_matrix.

DR InterPro; IPR003036; Gag_p30.

DR InterPro; IPR008916; Retrov_capsid_C.

DR InterPro; IPR001878; Znf_CCHC.

DR Pfam; PF01140; Gag_P30; 1.

DR Pfam; PF00098; Zf-CCHC; 1.

DR PRINTS; PR00939; C2HCZNFINGER.

DR SMART; SM00343; Znf_C2HC; 1.

DR PROSITE; PS0158; ZF_CCHC; 1.

KW Metal-binding; Zinc; Zinc-finger.

SQ SEQUENCE 524 AA; 59449 MW; 3825CCA687E1F450 CRC64;

Alignment Scores:

Pred. No.: 5,14e-166 Length: 524
 Score: 2774.00 Matches: 515
 Percent Similarity: 99.4% Conservative: 6
 Best Local Similarity: 98.3% Mismatches: 3
 Query Match: 97.4% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3_copy_585_2156 (1-1572) x Q4VPZ3_9GAMR (1-524)

QY 1 ATGGGACAGCGGTGACGCCCTCTAGTTTGAATCTCGACCAATTCGAGCTGAAGTTAAA 60
 DB 1 MetGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValLys 20
 QY 61 TCCAGGCTCATATTTGTCTAGTTTCAAGTAAAGGAGCCTTCGCGACACTTTCTGTGTC 120
 DB 21 SerArgAlaHisLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 40
 QY 121 TCTGAATGGCGGACATTCGATTTGGATGGCATCAGAGGGACCTTTAATCTGAGATT 180
 DB 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
 QY 181 ATCTGCTGCTTAAGCAGTTATTTTTCAGACTGACCGGCTCTCATCCGATCAGGAG 240
 DB 61 IleuAlaValLysAlaIleIlePheGlnThrGlyProGlySerHisProAsnGlnGlu 80
 QY 241 CCTTATATCTTACGTGGCAAGATTTCGACAGGATCTCCGCCATCGGTTAAACCATGG 300
 DB 81 ProTyrlleuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
 QY 301 CTGAATTAAGCAAGAAAGCCAGGTCCCGCAATTCCTGCTCTTGAGAGAGAAAAACAAC 360
 DB 101 LeuAsnLysProAspGlyProGlyProArgileLeuAlaLeuGlyGluLysAsnLysHis 120
 QY 361 TCGGCTGAAGAAAGTCAAGCCCTCTCATATCTACCCGAGATTGAGGAGCCACCGGCT 420
 DB 121 SerAlaGluLysValLysProSerProHisIleIleTrpProGluIleGluProProAla 140
 QY 421 TGGCCGGAACCCCAATCTGTTCCGCCACCCCTTATCTGACAGGTCGCGGAGGGGA 480
 DB 141 TrpProGluProGlnSerValProProProProTrpProTrpLeuAlaGlnGlyAlaAlaArgGly 160
 QY 481 CCCTTTGCCCTCTCTGGAGCTCCGGCGGTGGAGGACCTGCTCGAGGGACTCGAGCGCG 540
 DB 161 ProSerAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
 QY 541 AGGGGCGCCACCCCGAGCGGACGACGAGATTCGCGCATTACCGCTGCGCAGCTAGCGC 600
 DB 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrcGly 200

QY 601 CCTCCACACCGGGGGGCAATTGTCAGCCCTCCAGTATTGGCCCTTTTCTTCTCGAGAT 660
 DB 201 ProProThrProGlyGlyGlnLeuGlnProLeuGlnTyTrpProPheSerSerAlaAsp 220
 QY 661 CTCTATAATTGGAAACTAACCATCCCTCTTCGGAGGATCCCAACGGCTCACGGG 720
 DB 221 LeuTyraSerTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
 QY 721 TTGTGGAGTCCCTTATGTTCTCTCACAGCCTACTTTGGGATGATTGTCAACAGCTGTG 780
 DB 241 LeuValGlnSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
 QY 781 CAGACACTTTCACAAACCGAGGACGAGAGAAATCTTATTAGAGGCTAGAAAAAATGTT 840
 DB 261 GlnThrLeuPheThrThrGluArgGluArgileLeuLeuGluAlaArgLysAsnVal 280
 QY 841 CCTGGGCGCCACGGCGCACCCACCGCTTGCAAAATGAGATTGATGATGGGATTTCCCTTA 900
 DB 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
 QY 901 ACTCGCCCGGTTGGGACTACACACGCTGAAGGTAGGAGAGCTTGAATAATCTATCGC 960
 DB 301 ThrArgProGlyTrpAspTyraSerThrAlaGluGlyArgGluSerLeuLysIleTyArg 320
 QY 961 CAGGCTCTGTGGCGGCTCTCCGGGCGCTCAAGACGGCCCACTAATTTGGCTTAAGGTA 1020
 DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
 QY 1021 AGAGAAATGATGACGGGACCGGAATGAACCCCTCTCTTTTCTTTGAGAGGCTCTTGAA 1080
 DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluTrpLeuGlu 360
 QY 1081 GCCTTCAGGCGGTACACCCCTTTTGATCCACCTTCAGAGCCCAAGAGCTCAGTCGCT 1140
 DB 361 AlaPheArgArgTyThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
 QY 1141 TTGGCTTTTATAGGACAGTCCAGCTTCGATATTAGAAAGAGCTTCAGAGACTGGAAGG 1200
 DB 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
 QY 1201 TTACAGAGGCTCAGTTACGTGATCTAGTGAAGAGGACAGAGAAAGTATTATTACAAAGG 1260
 DB 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyTrpLysArg 420
 QY 1261 GAGACAGAAAGAAAGGGAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 DB 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluGluArgGluGluArg 440
 QY 1321 CGTAATAAACCGCAAGAGAGAAATTTGACTTAAGATCTTGGCTGCGAGTGGTTGAAGGAAA 1380
 DB 441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLys 460
 QY 1381 AGCAATACGAAAGAGAGAGAGATTTTAGAAAATTAGGTTCAGGCCCTAGACAGTCAGG 1440
 DB 461 SerAsnThrGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
 QY 1441 AACCTGGCCATAGGACCCCACTCGACAGGACCAATGTCATATTGTAAGAAGAGAGGA 1500
 DB 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrcLysGluLysGly 500
 QY 1501 CACTGGCAGGAAGTCCGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 DB 501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyLeuLysValLeuAlaLeuGlu 520
 QY 1561 GAAGATAAAGAT 1572
 DB 521 GluAspLysAsp 524

RESULT 5

Q6UUZ3_9GAMR

ID Q6UJZ3_9GAMR PRELIMINARY;

PRT; 524 AA.

AC Q6UJZ3;

DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)
DE Gag protein.
GN Name=gag;
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammaretrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX PubMed:14563152; DOI=10.1128/JVI.78.5.2502-2509.2004;
RA Scobie L., Taylor S., Wood J.C., Suling K.M., Quinn G., Meikle S.,
RA Patience C., Schuurman H.J., Onions D.E.;
RT "Absence of replication-competent human-tropic porcine endogenous
retroviruses in the germ line DNA of inbred miniature swine.";
RL J. Virol. 78:2502-2509(2004).
DR EMBL; AY368585; AAQ73934.1; -; Genomic DNA.
DR EMBL; AY368584; AAQ73936.1; -; Genomic DNA.
DR HSP; P03322; 1A6B.
DR SMR; Q6UJZ3; 7-98.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001878; ZnF_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; ZnF_C2HC; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 524 AA; 59474 MW; 72178B9541475920 CRC64;

Alignment Scores:
Pred. No.: 524
Score: 2757.00 Matches: 511
Percent Similarity: 99.0% Conservative: 8
Best Local Similarity: 97.5% Mismatches: 5
Query Match: 96.8% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x Q6UJZ3_9GAMR (1-524)
QY 1 ATGGGACAGCGTGCAGCACCCCTCTTAGTTGACTCTCGACCATTTGGACTGAAGTTAAA 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpIlePheGluVallys 20
QY 61 TCCAGGCTCATAAATTGTTCAGTTTAAAGAGGACCTTGGCAGACTTTCTGTGTC 120
DB 21 SerArgAlaHisAsnLeuSerValGlnVallysGlyProTrpGlnThrPheCysVal 40
QY 121 TCTGAATGGCCGACATTCGATGTGGATGGCCATCAGAGGGACCTTTAAATTCGAGATT 180
DB 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
QY 181 ATCTGGCTGTTAAAGCAGTTATTTTTCAGACTGACCGGCTCTCATCCGATCAGGAG 240
DB 61 IleLeuAlaVallysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
QY 241 CCCTATATCCTTACCTGGCAAGATTGGCAGAGGATCTCTCCGCCATGGGTTAAACCATGG 300
DB 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProProTrpVallysProTrp 100
QY 301 CTGAATAAGCCNAGAAAGCAGGTCCTCCGAATCTGGCTCTTGGGAGAGAAAACAAACAC 360
DB 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCCGGCT 420

DB 121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProAla 140
QY 421 TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGGTCCCGCAGGGGA 480
DB 141 TrpProGluProGlnSerValProProProTyrProAlaGlnGlyAlaAlaArgGly 160
QY 481 CCCTTTGCCCTCTCTGGAGCTCCGCGGTGGAGGACCTGCTGAGGAGACTCGAGACGG 540
DB 161 ProSerAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
QY 541 AGGGCGCCACCCCGAGCGGACAGATCCGACATTACCGCTCGCACCGTACGGC 600
DB 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly 200
QY 601 CCTCCACACCGGGGGCCAAATTCGAGCCCTCAGATATTGGCCCTTTCTTCTGCAGAT 660
DB 201 ProProIleProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220
QY 661 CTCTATAATTGGAACCTAACCATCCCTTTCTCGAGGATCCCAACGCTCACGGGG 720
DB 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACGAGCTACTTGGGATGATTGTCAACAGCTGCTG 780
DB 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
QY 781 CAGACACTTTCACACCGAGGAGCGAGAGAAATCTATTAGAGGCTAGAAAAAATGTT 840
DB 261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
QY 841 CCTGGGCGGACGCGGACCCACCGGTTGCAAAATGAGATTGATGATGGGATTTCCCTTA 900
DB 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
QY 901 ACTCGCCCGGTTGGGACTACAACACGCTGAAGCTAGGAGAGCTTGAAATCTATCGC 960
DB 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
QY 961 CAGCTCTGTGGCGGCTCTCCGGGCGCTCAAGACGCGCCCACTAATTTGGCTAAGTA 1020
DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
QY 1021 AGAGAGTGATGACGAGGACCGAATGAACCCCTCTGTTTTCTTTGAGAGCTCTTGAA 1080
DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
QY 1081 GCCTTCAGGCGGTACACCCCTTTTATCCCACTCAGAGGCGCCAAAAGCTCAGTGGCT 1140
DB 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
QY 1141 TTGGCCTTTATAGACAGTTCAGCCTTGGATATTAGAAGAAGCTTCAGAGACTGGAAGG 1200
DB 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
QY 1201 TTACAGGAGGCTGAGTTACGTGATCTAGTGAAGGAGGACAGAAAGTATATTACAAAAGG 1260
DB 401 LeuGlnGluAlaGluLeuArgAspLeuVallysGluAlaGluLysValTyrTyrLysArg 420
QY 1261 GAGACAGAAAGAAAGGAAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluGluGluGluGluGluArg 440
QY 1321 CGTAAATAACCGCAGAGAGAGAAATTTGAGTAAAGTATTTAGGTCCAGCCCTAGACAGT 1380
DB 441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLys 460
QY 1381 AGCAATACGGAAGAGAGAGAGATTTTAAAGAAATTTAGGTCCAGCCCTAGACAGTCA 1440
DB 461 SerAsnArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
QY 1441 AACCTGGGCAATAGGACCCCACTCGACAGGACCAATGTGCATATTGTAAAGAGAGAG 1500

Db 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
Qy 1501 CACTGGCGCAAGAACTGCCCAAGAGGAAACAAAGACCAAGGATCCTAGCTTAGAA 1560
Db 501 HisTrpAlaArgAspCysProLysLysGlyAsnLysGlyLeuLysValLeuAlaLeuGlu 520
Qy 1561 GAAGATAAAGAT 1572
Db 521 GluAspLysAsp 524

RESULT 6
ID Q6T3V4_9GAMR PRELIMINARY; PRT; 524 AA.
AC Q6T3V4;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Gag protein.
GN Name=gag;
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammatetrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Scobie L., Taylor S., Wood J.C., Suling K.M., Quinn G., Patience C.,
RA Schuurman H.-J., Onions D.E.;
RT "Absence of replication competent human-tropic porcine endogenous
RT retroviruses in the germ line DNA of inbred miniature swine.";
EL J. Virol. 78:0-0(2004).
DR EMBL; AY437841; AAR28102.1; -; Genomic_DNA.
DR HSSP; P03332; 1A6B.
DR SMR; Q6T3V4; 7-98.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_P30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_P30; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS50158; ZF_CCHC; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 524 AA; 59516 MW; 4B6B6C68FE930E32 CRC64;

Alignment Scores:
Pred. No.: 2,56e-164 Length: 524
Score: 2747.00 Matches: 509
Percent Similarity: 98.7% Conservative: 8
Best Local Similarity: 97.1% Mismatches: 7
Query Match: 96.5% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x Q6T3V4_9GAMR (1-524)

Qy 1 ATGGGACAGCGGTGACAGCCCTCTAGTTGTGACTCTCGACCAATTGGACTGAAGTTAAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValys 20
Qy 61 TCAGGCGCTCATATTTGTCAAGTTTCAGTTTAAAGAGGACCTTGGCAGACTTCTGTGTC 120
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 40
Qy 121 TCTGAATGGCGCATTCGATGTTGGATGGCATCAGAGGGGACCTTTAATTCGAGATT 180
Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluile 60
Qy 181 ATCTGGCTGTAAAGCAGTTATTTTTCAGACTGGACCCCGCTCTCATCCCGATCAGGAG 240

Db 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
Qy 241 CCCTATATCTTACGTGCAAGATTTCGCACAGGATCTCCCGCATGGTTAAACCATGG 300
Db 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
Qy 301 CTGAATAAGCCAAGAACCCAGGTCCCGCAATTCCTGGCTCTGGAGAGAAAAACAACAC 360
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
Qy 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGAGCCACCGCT 420
Db 121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProProAla 140
Qy 421 TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGCAGAGGTGCCCGAGGGA 480
Db 141 TrpProGluProGlnSerValProProProTyrProAlaGlnGlyAlaAlaArgGly 160
Qy 481 CCCTTTGCCCTCTCTGGAGCTCCCGCGGTGGAGGACCTGTCTGAGGAGCTCGGAGCCGG 540
Db 161 ProSerAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
Qy 541 AGGGCGCCACCCCGAGCGGACAGACGAGATCCGCACTTACCGCTGCGCACGTACGGC 600
Db 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly 200
Qy 601 CTTCCACACCGGGGGCCAAATTGCAGCCCTCCAGATTGGCCCTTTCTTCTTCGCAGAT 660
Db 201 ProProleProGlyGlyGlnLeuGlnProleuGlnTyrTrpProPheSerSerAlaAsp 220
Qy 661 CTCTATAATTGGAATAATAACATCCCTTTCTCGGAGGATCCCAACGCTCACCGGG 720
Db 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
Qy 721 TTGCTGAGTCCCTTATGTTCTCTCACAGCTACTTGGAGATGATTGTCAACAGCTGCTG 780
Db 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
Qy 781 CAGACACTCTTCAACCGGAGGACGAGAGAGTAATCTATTAGAGGCTAGAAAAAATGTT 840
Db 261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
Qy 841 CCTGGGCGCCAGCGGCGCACCGCTTGCAAAATGAGATTGACATGGGATTTCCCTTA 900
Db 281 ProGlyThrAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProleu 300
Qy 901 ACTGCCCCGGTTGGGACTACAACACGCTGAAGGTAGGAGAGCTTGAATAATCTATCGC 960
Db 301 IleArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
Qy 961 CAGGCTCTGCTGGGGGTCTCCGGGGCGCTCAAGAGCGGCCACTAATTTGGCTAAGGTA 1020
Db 321 GlnAlaLeuValAlaGlyLeuArgLysValSerArgArgProThrAsnLeuAlaLysVal 340
Qy 1021 AGAGAACTGATCAGGAGCACCAATGAACCCCTCTCTTTCTTTGAGAGGCTCTTGAA 1080
Db 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
Qy 1081 GCCTTCAGGCGGTACACCCCTTTTGATCCCACTCAGAGCGCCCAAAAGCTCTAGTGCT 1140
Db 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
Qy 1141 TTGCGCTTTTATAGACAGCTCAGCTTGATATTAGAAGAGCTTTCAGAGACTTGGAGGG 1200
Db 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGly 400
Qy 1201 TTACAGGAGCTCAGTTACGTGATCTAGTGAAGGAGGACAGAGAAAGTATATTACAAAAG 1260
Db 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420
Qy 1261 GAGACAGAAAGAAAGGGAACAAAGAAAAGAGAGAGAAAGAGAGGAAAGGAGGAAGA 1320

```
Db      421  GluThrGluGluGluArgGluGlnArgLysGluArgGluArgGluGluArg 440
Qy      1321 CGTAAATAACCGCAGAGAGAAATTTGACTAAGATCTTGCGCTGCGAGTGTGAAGGGAAA 1380
Db      441  ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaValValGluGlyLys 460
Qy      1381  AGCAATACGGAAGAGAGAGAGATTTTAGGAAATTTAGTCCAGGCCCTAGACAGTCAGGG 1440
Db      461  SerAsnArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
Qy      1441  AACCTGGGCAATAGGACCCCACTCGACAAGGACCAATGTGCATATTTGTAAGAAGAGAGGA 1500
Db      481  AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
Qy      1501  CACTGGGCAAGAACTGCCCCAAGAGGGAACCAAGGACCAAGATCCTAGCTCTAGAA 1560
Db      501  HisTrpAlaArgAspCysProLysLysGlyAsnLysGlyLeuLysValLeuAlaLeuGlu 520
Qy      1561  GAAGATAAAGAT 1572
Db      521  GluAspLysAsp 524
```

RESULT 7

```
Q8UMA0_9GAMR
ID      Q8UMA0_9GAMR PRELIMINARY; PRT; 524 AA.
AC
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Group specific antigen.
GN      Name=gag;
OS      Porcine endogenous retrovirus.
OC      Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC      Gamaretrovirus; 1-Mammalian type C virus group.
OX      NCBI_TaxID=61673;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=20219394; PubMed=10756014;
RX      DOI=10.1128/JVI.74.9.4028-4038.2000;
RA      Czaderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;
RT      "Establishment and characterization of molecular clones of porcine
RL      endogenous retroviruses replicating on human cells.";
RL      J. Virol. 74:4028-4038(2000).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=21851083; PubMed=11861838;
RX      DOI=10.1128/JVI.76.6.2714-2720.2002;
RA      Niebert M., Rogel-Galliard C., Chardon P., Toenjes R.R.;
RT      "Characterization of Chromosomally Assigned Replication-Competent Gamma
RT      Porcine Endogenous Retroviruses Derived from a Large White Pig and
RT      Expression in Human Cells.";
RL      J. Virol. 76:2714-2720(2002).
DR      EMBL; AJ279056; CAC82501.1; -; Genomic_DNA.
DR      HSSP; P03332; 1AGB.
DR      SMR; Q8UMA0; 7-98.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0046872; C:metal ion binding; IEA.
DR      GO; GO:0003676; F:nucleic acid binding; IEA.
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      GO; GO:0019068; P:viral assembly; IEA.
DR      InterPro; IPR003036; Gag_p30.
DR      InterPro; IPR000840; G_retro.matrix.
DR      Pfam; PF01140; Gag_MA; 1.
DR      Pfam; PF02093; Gag_p30; 1.
DR      Pfam; PF00098; zf_CCHC; 1.
DR      PRINTS; PR00939; C2HCZNFINGER.
DR      SMART; SM00343; ZNF_C2HC; 1.
DR      PROSITE; PS50158; ZF_CCHC; 1.
KW      Metal-binding; Zinc; Zinc-finger.
SQ      SEQUENCE 524 AA; 59502 MW; A5PD9F31FB3291AE CRC64;
```

Alignment Scores:

```
Pred. No.:      2.56e-164      Length:      524
Score:          2747.00      Matches:      508
Percent Similarity: 98.9%      Conservative: 10
Best Local Similarity: 96.9%      Mismatches: 6
Query Match:      96.5%      Indels:      0
DB:             2             Gaps:      0

US-10-723-552-3_copy_585_2156 (1-1572) x Q8UMA0_9GAMR (1-524)
Qy      1  ATGGAGCAGAGCGGTGAGACCCCTCTTAGTTTGTACTCTCGACCAATTGGACTGAAGTTAAA 60
Db      1  MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValLys 20
Qy      61  TCCAGGGCTCATATTTGTCTAGTTTCAGTTTAAAGAGGACCTTCGSCAGACTTCTGTCTC 120
Db      21  SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 40
Qy      121  TCTCAATGGCCGACATTCGATTTGGATGGCCATCAGAGGGGACCTTTAAATCTGAGATT 180
Db      41  SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
Qy      181  ATCTGCTGTGTTAAAGCAGATTATTTTCAGACTCGACCCGGCTCTCATCCCGATCAGGAG 240
Db      61  IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
Qy      241  CCCTATATCCTTACGTGCGCAAGATTTCGCGAGGATCCTCCGCCATGGTTAAACCATGG 300
Db      81  ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpTrpValLysProTrp 100
Qy      301  CTGAATAAGCCAAAGAACCCAGTCCCGAATTCGGCTCTTGGAGAGAAAAACAAACAC 360
Db      101  LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
Qy      361  TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGACCCACCGGCT 420
Db      121  SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProProAla 140
Qy      421  TGGCGGGAACCCCAATCTGTTCGCCCCACCCCTTATCTGGCCACAGGGTGC CGCAGGGGA 480
Db      141  TrpProGluProGlnSerValProProProTrpTrpProAlaGlnGlyAlaAlaArgGly 160
Qy      481  CCCTTTGCCCTCTCTGGAGCTCCGGCGGTGAGGAGACCTCTGCGAGGACTCGGAGCCGG 540
Db      161  ProSerAlaProProGlyAlaProAlaValGluGlyProValAlaGlyThrArgThrArg 180
Qy      541  AGGGGCGCCACCCCGAGCGGACAGAGAGATCCGCGACATTACCGCTCCGCGACGTACGGC 600
Db      181  ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly 200
Qy      601  CCTCCACACCGGGGGCCCAATTCAGCCCTCCAGTATTCGCCCTTTTCTTCTGCAGAT 660
Db      201  ProProIleProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220
Qy      661  CTCTATAATTGAAAACTAACCATCCCCCTTTCTCGGAGGATCCCCAACGCGCTCACGGGG 720
Db      221  LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
Qy      721  TTGGTGGAGTCCCTTATGTTCTCTCACAGCCTACTTGGGATGATTGTCAACAGCTGCTG 780
Db      241  LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
Qy      781  CAGACACTCTTTCACACCGGAGCGAGAGAGAATCTATTAGAGGCTAGAAAAAATGTT 840
Db      261  GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
Qy      841  CCTGGGCGGACGGCGGACCCACCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTAA 900
Db      281  ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
Qy      901  ACTGCCCGCGTTGGGACTACAACACGGCTGAAGTAGGGAGAGCTTGAAAAATCTATGCG 960
Db      301  ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
```

QY 961 CAGGCTCTGGTGGGGCTCTCGGGCGCCTCAAGACGGCCCACTAATTTGGCTAAGGTA 1020
 Db GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgProThrAsnLeuAlaLysVal 340
 QY 1021 AGAGAAGTGATGACGGGACCGAATGAACCCCTCTGTTTCTTGGAGAGGCTCTTGAA 1080
 Db ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
 QY 1081 GCCTTCAGCGGTTACACCCCTTTTGTATCCCTCAGAGGCGCCCAAAAGCCCTCAGTGCT 1140
 Db AlaPheArgGlyPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
 QY 1141 TTGGCCCTTATAGGACAGCTAGCCTTGATATTAGAAAGAGCTTCAGAGACTGAAGGG 1200
 Db LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
 QY 1201 TTACAGAGGCTGAGTCTACGTGATCTAGTGAAGGAGGAGAGAAAGTATATTACAAAGG 1260
 Db LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrLysArg 420
 QY 1261 GAGACAGAAAGAAAGGAAACAAAGAAAGAGAGAGAAAGAGAGGAGGAGGAAAGA 1320
 Db GluThrGluGluGluArgAspGlnArgLysGluArgGluArgGluArgGluArg 440
 QY 1321 CGTAAATAAACGGCAAGAGAAATTTGACTAAGATCTTGGCTGAGTGTTGAAGGAAA 1380
 Db ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaValGluGlyLys 460
 QY 1381 AGCAATACGGAAGAGAGAGATTTAGGAATATTAGGTACAGCCCTAGACACTCAGG 1440
 Db SerAsnArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
 QY 1441 AACCTGGGCAATAGGACCCCACTCGCAAGACCAATGTGCATATTGTAAGAAAGAGGA 1500
 Db AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
 QY 1501 CACTGGGCAAGAACTGCCCAAGAGGAAACAAAGACCAAGGATCTAGCTCTAGAA 1560
 Db HisTrpAlaArgAspCysProLysLysGlyAsnLysGlyLeuLysValLeuAlaLeuGlu 520
 QY 1561 GAAGATAAAGAT 1572
 Db GluAspLysAsp 524

RESULT 8

Q90RM0_9GAMR
 ID Q90RM0_9GAMR PRELIMINARY; PRT; 524 AA.
 AC Q90RM0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Group specific antigen.
 GN Name: gag;
 OS Porcine endogenous type C retrovirus.
 OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Gammatetrovirus.
 OX NCBI_TaxID=70540;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21256017; PubMed=11356953;
 RX DOI=10.1128/JVI.75.12.5465-5472.2001;
 RA Krach U., Fischer N., Czauderna F., Toenjes R.R.;
 RT "Comparison of replication-competent molecular clones of porcine
 endogenous retrovirus class A and class B derived from pig and human
 cells.";
 RT J. Virol. 75:5465-5472 (2001).
 RL EMBL; AJ293656; CAC39616.1; -; Genomic_DNA.
 DR HSSP; P03332; 1A6B.
 DR SMR; Q90RM0; 7-98.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.

GO; GO:0019068; P:viral assembly; IEA.
 DR InterPro; IPR003036; Gag_p30.
 DR InterPro; IPR000840; G retro matrix.
 DR InterPro; IPR001878; ZnF_CCHC.
 DR Pfam; PF01140; Gag_MA; 1.
 DR Pfam; PF02093; Gag_p30; 1.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PROSITE; PS00343; ZnF_C2HC; 1.
 DR PROSITE; PS0158; ZnF_CCHC; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 524 AA; 59486 MW; 83B73BB3405C9D8A CRC64;

Alignment Scores:

Pred. No.: 3,42e-164 Length: 524
 Score: 2745.00 Matches: 508
 Percent Similarity: 98.9% Conservative: 10
 Best Local Similarity: 96.9% Mismatches: 6
 Query Match: 96.4% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x Q90RM0_9GAMR (1-524)

QY 1 ATGGGACAGACGGTGAAGACCCCTCTTAGTTTGACTCTCGACCATTTGGACTGAAGTTAAA 60
 Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValLys 20
 QY 61 TCCAGGGCTCATATTTGTCAAGTTTCAGTTTAAAGAGGACCTTCGCAGACTTCTGTGTC 120
 Db 21 SerArgAlaHisAsnLeuSerValGlnValLysGlyProTrpGlnThrPheCysVal 40
 QY 121 TCTGAATGGCCGACATTCGATTTGGATGGCCATCAGAGGGGACCTTTAATTTCTGAGATT 180
 Db 41 SerGluTrpTrpThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
 QY 181 ATCTGGCTGTAAAGCAGTTATTTTCAGACTGGACCCGGCTCTCATCTCCGATCAGAG 240
 Db 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
 QY 241 CCTTATATCTTACGTGCGCAAGATTTGCGAGAGGATCTCCGACATGGTGTAAACCATGG 300
 Db 81 ProTrpIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
 QY 301 CTGAATAAGCAAGAAAGCCAGGTCCCGGAATTCCTGCTCTTGAGAGAGAAAAACAAACAC 360
 Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
 QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGACCCAGCGCT 420
 Db 121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProAla 140
 QY 421 TGGCCGGAACCCCAATCTGTTCCCAACCCCTTATCTGGCACAGGTTGCCGCGAGGGA 480
 Db 141 TrpProGluProGlnSerValProProProTrpProAlaGlnGlyAlaAlaArgGly 160
 QY 481 CCTTTTCCCTCTCTGGAGCTCCGGCGTGGAGGACCTCTGCGAGGACTCGGAGCGGG 540
 Db 161 ProSerAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
 QY 541 AGGGGCGCCACCCCGGAGCGACAGACGAGATTCGACATTACCGCTGCGCACGTACGGC 600
 Db 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly 200
 QY 601 CTCTCCACACCGGGGGCGAATTGACAGCCCTCTCAGTATTGGCCCTTTCTTCTGAGAT 660
 Db 201 ProProIleProGlyGlnLeuGlnProLeuGlnTrpProPheSerSerAlaAsp 220
 QY 661 CTCTATAATTGGAAAACTAACCACTCCCTCTTCTCGGAGGATCCCCAACCCCTCACCGGG 720
 Db 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnAlaGluLeuThrGly 240
 QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACCGCTTCTTGGGATGATTCTCAACAGCTCTGT 780

```
Db      241  LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
Qy      781  CAGACACTTTCACACACCGAGGAGGAGAGAAATCTATTAGAGCTAGAAAAATGTT 840
Db      261  GlnThrLeuPheThrThrGluGluArgLysArgIleLeuLeuGluAlaAArgLysAsnVal 280
Qy      841  CCTGGGCGGACGGCGGACCCACCGGGTTGC AAAATGAGATTGACATGGGATTTCCCTTA 900
Db      281  ProGlyAlaAspGlyArgProThrGlnLeuGlnGlnIleAspMetGlyPheProLeu 300
Qy      901  ACTGCCCGCGTGTGGACTACACACACGCTGAAGGTAGGAGAGCTTGAAAATCTATCGC 960
Db      301  ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
Qy      961  CAGCTCTGGTGGCGGGTCTCGGGGCGCCTCAGACGGCCCACTAAATTTGGCTAAGGTA 1020
Db      321  GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
Qy      1021  AGAGAAGTGATGCAGGAGCCGAATGAACCCCTCTGTTTCTTTGAGAGCTCTTGGA 1080
Db      341  ArgGluValMetGlnGlyProAsnGluProSerValPheLeuGluArgLeuMetGlu 360
Qy      1081  GCCTTCAGCGGTACACCCCTTTTGTATCCCACTCAGAGGCCCAAAAAGCCTCAGTGGCT 1140
Db      361  AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerMetAla 380
Qy      1141  TTGGCTTTATAGACAGCTAGCTTGATATATAGAAAGCTTCAGAGCTGGAAGGG 1200
Db      381  LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
Qy      1201  TTACAGAGGCTGAGTTACGTGATCTAGTGAAGGCGCAGAGAAAGTATATTACAAAGG 1260
Db      401  LeuGlnGluAlaGluLeuHisAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420
Qy      1261  GAGACAGAAGAAAGGAAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db      421  GluThrGluGluGluArgGluGlnArgLysGluArgGluArgGluGluArgGluArg 440
Qy      1321  CGTAATAACCGCAAGAGAGAAATTTGACTAAGATCTTGGCTGCGAGCTGTTGAAGGGAAA 1380
Db      441  ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValAlaGluGlyLys 460
Qy      1381  AGCAATACGGAAGAGAGAGAGATTTTAGGAAATTTAGTCCAGGCCCTTAGACAGTCAGGG 1440
Db      461  SerAsnArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
Qy      1441  AACCTGGGCAATAGGACCCCACTCGACAAGGACCAATGTGCATATTGTAAAGAAAGAGA 1500
Db      481  AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
Qy      1501  CACTGGGCAAGAACTGCCCCAAGAGGAAACAAAGGACCAAGAGTCTTAGCTCTAGAA 1560
Db      501  HisTrpAlaArgAspCysProLysLysGlyAsnLysGlyLeuLysValLeuAlaLeuGlu 520
Qy      1561  GAAGTAAAGAT 1572
Db      521  GluAspLysAsp 524

RESULT 9
Q6UJZ5_9GAMR
ID Q6UJZ5_9GAMR PRELIMINARY; PRT; 524 AA.
AC Q6UJZ5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gag protein.
GN Name=gag;
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammatetrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RX PubMed=14963152; DOI=10.1128/JVI.78.5.2502-2509.2004;
RA Scobie L., Taylor S., Wood J.C., Suling K.M., Quinn G., Meikle S.,
RA Patience C., Schuurman H.J., Onions D.E.;
RT "Absence of replication-competent human-tropic porcine endogenous
RT retroviruses in the germ line DNA of inbred miniature swine.";
RL J. Virol. 78:2502-2509(2004).
DR EMBL; AY368583; AAQ73933.1; -; Genomic_DNA.
DR HSSP; P03332; 146B.
DR SMR; Q6UJZ5; 7-98.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_P30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_P30; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS0158; ZF_CCHC; 1.
DR Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 524 AA; 59525 MW; B2AF9A672406B8EC CRC64;

Alignment Scores:
Pred. No.: 6,19e-163 Length: 524
Score: 2725.00 Matches: 507
Percent Similarity: 98.5% Conservative: 9
Best Local Similarity: 98.8% Mismatches: 8
Query Match: 95.7% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x Q6UJZ5_9GAMR (1-524)
Qy 1 ATGGGACAGACGGTGACGACCCCTCTAGTTTGACTCTCGACCATTGAGACTGAAGTTAAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValLys 20
Qy 61 TCCAGGGCTCATAAATTTGTCAAGTTTCAGTTTAAGAGGACCTTGGCAGACTTTCTGTGTC 120
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 40
Qy 121 TCTGAATAGCCGCGACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAAATCTGAGATT 180
Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
Qy 181 ATCTCTGGCTGTTAAGCAGATTATTTTTCAGACTCGACCCGGCTCTCATCCCGATCAGGAG 240
Db 61 IleuAlaValLysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
Qy 241 CCCTATATCTTACGTGCAAGATTGTCGACAGGATCTCCCGCCATGGGTTTAAACCATGG 300
Db 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
Qy 301 CTGAATAGCCAAAGAACCCAGGTCCCGAATCTGGCTCTTGGGAGAGAAAACAAACAC 360
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
Qy 361 TCGGCTGAAGAAGTCAGGCCCTCTCCTCATATCTACCCCGAGATTGAGAGCCACCGGCT 420
Db 121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProAla 140
Qy 421 TGGCGGGAACCCCAATCTGTTCCCGCCACCCCTTATCTGGCACAGGGTGC CGCAGGGGA 480
Db 141 CysProGluProGlnSerValProProProTyrProAlaGlnGlyAlaAlaArgGly 160
Qy 481 CCCTTTGCCCTCTCTGGAGCTCCGCGGTGAGGAGACCTGTCTGAGGAGCTCGGAGCCGG 540
Db 161 ProSerAlaProProGlyAlaProAlaValGluGluProAlaAlaGlyThrArgSerArg 180
Qy 541 AGGGCGGCCACCCCGGAGCGGACAGAGATCGGACATTACCGCTCGCACGTACGGC 600
```

Db 181 ArgGlyAlaThrProGluArgThrAspGluLeuAlaThrLeuProLeuArgThrTyrgly 200
Qy 601 CTCTCCACACCGGGGGGCAATTTGAGCCCTCCAGTATTGGCCCTTTTCTTCTGACAGT 660
Db 201 ProProIleProGlyGlyGlnLeuGlnProLeuGlnTyrrProPheSerSerAlaAsp 220
Qy 661 CTCTATAATTGAAATAACCATCCCTTTCTCGAGAGATCCCAACGCTCACGCGG 720
Db 221 LeuTyrrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
Qy 721 TTGTGTAGTCCCTTATCTCTCCACAGCTACTTGGGATGATGTCAACAGCTGCTG 780
Db 241 LeuMetGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
Qy 781 CAGACACTCTTCCACAAACGAGGAGCGAGAGAGATTTCTATTAGAGCTTAGAAAAATGTT 840
Db 261 GlnThrLeuPheThrGluGluArgGluArgGluLeuGluAlaArgLysAsnVal 280
Qy 841 CCTGGGCGGACGGGCGACCCACGCGGTGCAAAATAGATGATGATGGGATTTCCCTTA 900
Db 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
Qy 901 ACTGCGCCCGCTTGGGACTCAACACGCTGAGGTAGGAGAGCTTGAATCTATCGC 960
Db 301 ThrArgProGlyTrpAspTyrrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrrArg 320
Qy 961 CAGGCTCTGTGGGGGTCTCCGGGCGCTCAAGACGGCCCACTAATTTGGCTTAAGTA 1020
Db 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
Qy 1021 AGAAGATGATGAGGAGACCAATGAACCCCTCTGTTTTCTTGAGAGCTCTTGAA 1080
Db 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
Qy 1081 GCCTTCAGGCGGTACACCCCTTTGATCCACCTCAGAGGCCCAAAAGCCCTCAGTGCT 1140
Db 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
Qy 1141 TTGGCTTTATAGACAGTCAAGCTTGGATATTAGAAAGAGCTTCAGAGACTGGAAGG 1200
Db 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
Qy 1201 TTACAGGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrrLysArg 420
Qy 1261 GAGACAG 1320
Db 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluArgGluArgGluArgGlu 440
Qy 1321 CGTAAATAACCGCAAGAGAGAAATTTGACTAAGATCTTGGCTGAGTGGTTGAAGGAAA 1380
Db 441 ArgAsnLysArgGlnGlnLysAsnLeuThrLysIleLeuAlaValGluGlyLys 460
Qy 1381 AGCAATACGGAAGAGAGAGAGATTTTAGGAAATTTAGGTCAGGCCCTTAGACAGTCA 1440
Db 461 SerAsnArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
Qy 1441 AACTGGGCAATAGAGCCCACTCGACAGGACCAATGTGCATATTGTAAGAAGAGGA 1500
Db 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysLysGluLysGly 500
Qy 1501 CACTGGGCAAGGAACTGCCCCAAGAGGAAACAAAGGACCAAGGACCTTAGCTCTAGAA 1560
Db 501 HisTrpAlaArgAspCysProLysLysGlyAsnLysGlyLeuLysGluIleAlaLeuGlu 520
Qy 1561 GAAGATAAAGAT 1572
Db 521 GluAspLysAsp 524

RESULT 10

Q6UJZ2_9GAMR

ID Q6UJZ2_9GAMR PRELIMINARY; PRT; 524 AA.
AC Q6UJZ2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gag protein.
GN Name=gag;
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammaretrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14963152; DOI=10.1128/JVI.78.5.2502-2509.2004;
RA Scobie L., Taylor S., Wood J.C., Suling K.M., Quinn G., Meikle S.,
RA Patience C., Schuurman H.J., Onions D.E.;
RT "Absence of replication-competent human-tropic porcine endogenous
RT retroviruses in the germ line DNA of inbred miniature swine.";
RL J. Virol. 78:2502-2509(2004).
DR EMBL; AY368586; AAQ73935.1; -; Genomic_DNA.
DR HSSP; P03332; 1A6B.
DR SMR; O6UJZ2; 7-98.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_P30.
DR InterPro; IPR000840; G_Retro_matrix.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF011140; Gag_MA; 1.
DR Pfam; PF02093; Gag_P30; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PSS0158; zf_CCHC; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 524 AA; 59349 MW; 1F7CDC5A5B8F13F8 CRC64;

Alignment Scores:
Pred. No.: 7,15e-163 Length: 524
Score: 2724.00 Matches: 507
Percent Similarity: 98.1% Conservative: 7
Best Local Similarity: 96.8% Mismatches: 10
Query Match: 95.6% Indels: 0
DB: Gaps: 2

US-10-723-552-3_COPY_585_2156 (1-1572) x Q6UJZ2_9GAMR (1-524)

Qy 1 ATGGGACAGACGGTGCAGCCCTCTTGTAGTTTGAATCTCGACCATTCGACTGAAGTTAAA 60
Db 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValys 20
Qy 61 TCACGGCTCATTAATTTGTTCAGTTTCAGTTTAAGAGGACCTTGGCAGACTTCTGTGTC 120
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 40
Qy 121 TCTGAATGGCGGACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAATCTGAGATT 180
Db 41 SerGlnTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
Qy 181 ATCTGCTGCTTAAAGCAGTTATTTTTCAGACTGGACCCCGCTCTCATCCGATCAGAG 240
Db 61 IleLeuAlaValLysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
Qy 241 CCTATATCTTACGTGGCAAGATTGGCGAGAGATCTCCCGCATCGGTTAAACCATGG 300
Db 81 ProTyrrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
Qy 301 CTGAATTAAGCAAGAGCCAGGTCCCGCAATTTCTGGCTTTCGGAGAGAGAAACACAC 360
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120

Qy 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCACCGGCT 420
Db |||||
121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProProAla 140
Qy 421 TGGCGGGAACCCCAATCTGTTCCGCCACCCCTTATCTGGCACAGGGTGC CGCAGGGGA 480
Db |||||
141 GlyProGluProGlnSerValProProProTyrProAlaGlnGlyAlaAlaArgGly 160
Qy 481 CCCTTTGGCCCTCTCGAGCTCGCGCGTGGAGGACCTGCTGCAGGAGCTCGAGCGG 540
Db |||||
161 ProSerAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
Qy 541 AGGGCGCCACCCCGGAGCGACAGCAGATCGACATTACGCTCGCGCACGTACGGC 600
Db |||||
181 ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly 200
Qy 601 CCTCCACACCGGGGGGCAATGTGAGCCCTCCAGTATTCGCCCTTTCTTCGCGAT 660
Db |||||
201 ProProIleProGlyGlnLeuGlnProLeuGlnTyrTyrProPheSerSerAlaAsp 220
Qy 661 CTCTATAATGGAAACTAACATCCCTCTTCGAGAGATCCCAACGCTCAGCGGG 720
Db |||||
221 LeuTyrAsnTyrLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
Qy 721 TTGGTGGAGTCCCTTATGTTCTCTCACCGACTACTTCGGATGATTGTCAACAGCTGCTG 780
Db |||||
241 LeuValGluSerLeuMetPheSerHisGlnProThrTyrAspAspCysGlnGlnLeuLeu 260
Qy 781 CAGACACTCTTACACACCGAGGAGCGAGAGAAATCTTATTAGAGGCTAGAAAAATGTT 840
Db |||||
261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
Qy 841 CCTGGGCGGAGCGGCGGACCCACGCGGTGCAAAATGAGATTGACATGGGATTCCTCCTTA 900
Db |||||
281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
Qy 901 ACTGCGCCGGTGGGACTACAAACGCTCAAGGTAGGAGAGCTTCAAAATCTATCGC 960
Db |||||
301 ThrArgProGlyTyrAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
Qy 961 CAGGCTCTGGTGGCGGGTCTCGGGGCGCTCAAGACGCGCCACTAAATTTGGCTAAGGTA 1020
Db |||||
321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
Qy 1021 AGAGAAGTGTGACGAGGACCGAATGAACCCCGCTCTGTTTCTTGACAGGCTCTTGAA 1080
Db |||||
341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
Qy 1081 GCCTTCAGCGGTACACCCCTTTTGCATCCACCTCAGAGGCCCAAAAGCCTCAGTGGCT 1140
Db |||||
361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
Qy 1141 TTGGCCCTTTATAGGACAGTCAAGCTTGGATATTAGAAAAGCTTTCAGAGCTGGAAGG 1200
Db |||||
381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
Qy 1201 TTACAGAGGCTGAGTTAGTGTACTAGTAAAGGAGGAGAAAGTATATTACAAAGG 1260
Db |||||
401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420
Qy 1261 GAGACAGAGAAGAAAGGGAAACAAAGAAAAGAGAGAGAGAAAGGAGGAGGAGAAAGA 1320
Db |||||
421 GluThrGluGluGluArgAspGlnArgLysGluArgGluArgGluArgGluGluArg 440
Qy 1321 CGTATAAACCGCAAGAGAGAAGAAATTTGACTTTGGCTGCGTGTGCTGAGGGGAAA 1380
Db |||||
441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValAlaGluGlyLys 460
Qy 1381 AGCAATAACGGAAGAGAGAGAGATTTTAGGAAAATTTAGTCTGCGCCCTAGACAGTCAGG 1440
Db |||||
461 SerAsnArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
Qy 1441 AACCTGGGCAATAGGACCCCTCTCGACAAGGAGCAATGTGCATATTGTGTAAGAAAGAGGA 1500

Db 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
Qy 1501 CACTGGCGAAGAACTCCGCCCAAGAGGAAACAAAGACCAAGATCCTAGCTCTAGAA 1560
Db |||||
501 HisTrpAlaArgAspCysProLysGlyAsnLeuGlyThrGluAsnLeuThrLeuGlu 520
Qy 1561 GAACATAAAGAT 1572
Db |||||
521 GluAspLysAsp 524
RESULT 11
Q9NLT7_PIG PRELIMINARY; PRT; 664 AA.
ID Q9NLT7_PIG PRELIMINARY; PRT; 664 AA.
AC Q9NLT7_PIG PRELIMINARY; PRT; 664 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gag-pol (Fragment).
OS Sus scrofa domestica (domestic pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9825;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=Breid Duroc x German Landbreed;
RA Blusch J.H., Seelmeier S., von der Helm K.;
RT "Evidence for extreme sequence homogeneity of the PERV A/B 5'
LTR/leader/gag/prot region by predicted targeted amplification from a
Du x DL pig.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF147808; AAF65926.1; -; Genomic_DNA.
DR HSSP; Q9WJF4; 1BW4.
DR SMR; Q9NLT7; 7-38.
DR MEROPS; A02.020; -.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019068; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS0158; zf_CCHC; 1.
FT CHAIN 1 524 gag protein.
FT CHAIN 525 >664 aspartic protease.
FT NON_TER 664
SQ SEQUENCE 664 AA; 75259 MW; 188FA83D39FE2F02 CRC64;
Alignment Scores:
Pred. No.: 1,15e-161 Length: 664
Score: 2705.00 Matches: 497
Percent Similarity: 97.9% Conservative: 16
Best Local Similarity: 94.8% Mismatches: 11
Query Match: 95.0% Indels: 0
DB: 2 Gaps: 0
US-10-723-552-3_copy_585_2156 (1-1572) x Q9NLT7_PIG (1-664)
Qy 1 ATGGACAGACGGTGACGACCCCTCTTAGTTTGTACTCTCGACCATTTGAGCTGAAGTTAAA 60
Db |||||
1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20

QY 61 TCCAGGGCTCATATTTCTCAGTTTCAGGTTTAAAGAGGACCTTGGCAGACTTTCGTGTC 120
DB 21 SerArgAlaHisAsnLeuSerValGlnValIysGlyProTrpGlnThrPheCysAla 40
QY 121 TCTGAATGGCGGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAAATCTCAGATT 180
DB 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
QY 181 ATCTGGCTGTAAAGCAGTTTATTTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
DB 61 IleLeuAlaValIysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
QY 241 CCCTATATCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATGGGTTAAACCATGG 300
DB 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValIysProTrp 100
QY 301 CTGAATTAAGCCAAAGACCCAGGTCCTCCATATCTACCCCGAGATTGAGGAGCCACCGCT 360
DB 101 LeuAsnIysProArgIysProGlyProArgIleLeuAlaLeuGlyGluIysAsnLysHis 120
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCCTCATATCTACCCCGAGATTGAGGAGCCACCGCT 420
DB 121 SerAlaGluIysValGluProSerProArgIleTyrProGluIleGluGluProProThr 140
QY 421 TGGCCGGAACCCCAATCTGTTCCCTCCACCCCTTATCTGGCACAAGGTTGCCGAGGGGA 480
DB 141 TrpProGluProGlnProValProProProProTyrProAlaGlnGlyAlaValArgGly 160
QY 481 CCCTTTGCCCTCTCTGAGCTCCGGCGGTGGAGGACCTGCTGAGGAGCTCGAGCGGG 540
DB 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
QY 541 AGGGGCCCCACCCCGAGCGGACAGACGAGATCCGACATTACCGCTGCGCAGCTACGGC 600
DB 181 ArgGlyAlaThrProGluIysGlnLeuGlnProLeuGlnIleAlaLeuProLeuArgThrTyrGly 200
QY 601 CCTCCACACCGGGGGCCCAATTGCAGCCCTCCAGTATTGGCCCTTTCTTCTGCAAGAT 660
DB 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnIleTyrTrpProPheSerSerAlaAsp 220
QY 661 CTCTATAATTTGGAATACTAACCATCCCTCTTCTCGGAGGATCCCAACCGCTCACGGGG 720
DB 221 LeuTyrAsnTrpIysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACCAGCTACTTGGGATGATCTCAACAGCTCTG 780
DB 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
QY 781 CAGACACTCTTCAACCCGAGGAGCGAGAGAAATCTATTAGAGGCTAGAAAAAATGTT 840
DB 261 GlnThrLeuPheThrThrGluArgGluArgIleLeuLeuGluAlaArgIysAsnVal 280
QY 841 CTGGGGCCGACGGGCGGACCCAGCGGTTGCAAAATGAGATTGACATGGGATTCCTTTA 900
DB 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
QY 901 ACTCGCCCGGTTGGGACTACACACCGCTGAAGGTAGGAGGACCTTGAATCTATCCG 960
DB 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuIysIleTyrArg 320
QY 961 CAGGCTCTGGTGGCGGTCTCCGGGGCGCTCAAGACGGCCCACTAATTTGGCTTAAGGTA 1020
DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaIysVal 340
QY 1021 AGAAGATGTAGCAGGACCGAATGAACCCCTCTCTTTTCTTGAGAGGCTCTTGAA 1080
DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
QY 1081 GCCTTCAGCGGTACACCCCTTTTGATCCACCTCAGAGGCCCAAAAGCCTCAGTGGCT 1140
DB 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380

QY 1141 TTGGCCCTTTATAGACAGCTCAGCTTGGATATTAGAAGAAGCTTCAGAGACTGGAAGG 1200
DB 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgIysLysLeuGlnArgLeuGluGly 400
QY 1201 TTACAGAGGCTGAGTTACGTGATCTAGTGAAGGCGCAGAGAAAGTATATATACAAAGG 1260
DB 401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 420
QY 1261 GAGACAGAAGAAGGAGGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 421 GluThrGluGluLysGluGlnArgIysGluIysGluArgGluGluArgGluGluArg 440
QY 1321 CGTAATAACCGCAAGAGAAATTTGACTAAGATCTTGGCTGAGTGGTGTAAAGGAAA 1380
DB 441 ArgAspArgArgGlnGluLysAsnLeuThrIysIleLeuAlaValValGluGlyLys 460
QY 1381 AGCAATACGGAAGAGAGAGAGATTTTAGGAAATTTAGTTCAGGCCCTTAGACAGTCAGG 1440
DB 461 SerSerArgGluArgGluArgGluArgPheArgIysIleArgSerGlyProArgGlnSerGly 480
QY 1441 AACCTGGGCAATAGGACCCCACTCGACAAGGACCAATGTGCATATTGTAAAGAGAGAGA 1500
DB 481 AsnLeuGlyAsnArgThrProLeuAspIysAspGlnCysAlaTyrCysLysGluIysGly 500
QY 1501 CACTGGCAAGGAACTCCCAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
QY 1561 GAAGATTAAGAT 1572
DB 521 GluAspLysAsp 524
RESULT 12
Q8J4V8_9GAMR
ID Q8J4V8_9GAMR PRELIMINARY; PRT; 1718 AA.
AC Q8J4V8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gag-pol polyprotein.
OS Porcine endogenous retrovirus A.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=194958;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22172947; PubMed=12185278;
RA Bartosch B., Weiss R.A., Takeuchi Y.;
RT "PCR-based cloning and immunocytological titration of infectious
RL J. Gen. Virol. 83:2231-2240(2002).
DR EMBL; AY099323; AAM29192.1; -; Genomic_DNA.
DR HSP; P03355; I160.
DR SMR; Q8J4V8; 7-98.
DR MEROPS; A02.020; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G retro matrix.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR001989; Pept_Asp_AS.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.

DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF01140; Gag_MA; 1_CCHC.
 DR Pfam; PF02093; Gag_P30; 1.
 DR Pfam; PF00075; RnaseH; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT; 1; 1.
 DR Pfam; PF00098; zf_CCHC; 1.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS0879; RNASE H; 1.
 DR PROSITE; PS0158; zf_CCHC; 1.
 KW Polyprotein.
 SQ SEQUENCE 1718 AA; 193115 MW; 4E86215286C59246 CRC64;

Alignment Scores:
 Pred. No.: 1,266-161 Length: 1718
 Score: 2705.00 Matches: 497
 Percent Similarity: 97.9% Conservative: 16
 Best Local Similarity: 94.8% Mismatches: 11
 Query Match: 95.0% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x Q8J4V8_9GAMR (1-1718)

QY	1	ATGGGACAGCGGTGACGACCCCTCTAGTTGACTCTCGACCATGGACCTGAAGTTAAA	60
DB	1	MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg	20
QY	61	TCCAGGGCTCATATTTCTGCTAGTTTCAGGTTAAGAGGACCTTGGCAGACTTTCGTGTC	120
DB	21	SerArgAlaHisAsnLeuSerValGlnValIysIysGlyProTrpGlnThrPheCysAla	40
QY	121	TCTGAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAAATCTGAGATT	180
DB	41	SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle	60
QY	181	ATCTGGCTGTAAAGCAGTATTTTTCAGACTGGACCCGGCTCATCCCGATCAGGAG	240
DB	61	IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu	80
QY	241	CCCTATATCTTACCTGGCAAGATTGGCAGAGATCTCCGCCATGGTGTAAACCATGG	300
DB	81	ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValIysProTrp	100
QY	301	CTGAATAGCCAAAGCCAGGTCCTCCGAAATCTGGCTCTTGGAGAGAAAACAAACAC	360
DB	101	LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis	120
QY	361	TCGGCTGAAAAGTCAGCCCTCTCTCATATCTACCCCGAGATTGAGGACCCAGGCT	420
DB	121	SerAlaGluLysValGluProSerProArgIleTyrProGluIleGluGluProProThr	140
QY	421	TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGGTGGCCGAGGGGA	480
DB	141	TrpProGluProGlnProValProProProTyrProAlaGlnGlyAlaValArgGly	160
QY	481	CCCTTTGCCCCCTCTCTGGAGCTCCGCGGTGGAGGACCTGTGCGAGGACTCGGAGCCGG	540
DB	161	ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg	180
QY	541	AGGGCGGCCACCCCGGAGCGGACGAGATCGGCACATTACGCTCGGCACGTACGGC	600
DB	181	ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly	200
QY	601	CCTCCACACCGGGGGCCAAATTCAGCCCTCCAGTATTCGCCCTTTTCTTCGCAGAT	660
DB	201	ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp	220
QY	661	CTCTATAATGGAAAACCAATCCATCCCTTTCTCGAGGATGCCCAACGCTCACGGGG	720
DB	221	LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly	240

QY	721	TTGGTGGAGTCCCTTATGTTCTCTCACAGCCTACTTGGGATGATTGTCAACAGCTGCTG	780
DB	241	LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspPheCysGlnGlnLeuLeu	260
QY	781	CAGACACTCTTTCACACCGGAGCGAGAGAGAATTTCTATTAGAGGCTAGAAAAATGTT	840
DB	261	GlnThrLeuPheThrThrGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal	280
QY	841	CCTGGGCGGACGGCGGACCCACCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTA	900
DB	281	ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu	300
QY	901	ACTGCGCCCGTTCGGGACTACAACACGCTCAAGCTAGGAGAGCTTCAAAATCTATCGC	960
DB	301	ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg	320
QY	961	CAGCTCTGTGTGGCGGTCTCCGGGCGCTCAAGACGGCCCACTAAATTTGGCTAAGTA	1020
DB	321	GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal	340
QY	1021	AGAGAATGATGCAGGACCGAATGAACCCCTCTGTTTTTCTTGAGAGCTCTTGAA	1080
DB	341	ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu	360
QY	1081	GCCTTCAGGGGTACACCCCTTTTGTATCCACCTCAGAGGCCCAAAAGCCTCAGTGGCT	1140
DB	361	AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla	380
QY	1141	TTGGCCTTTATPAGACAGTCAGCCTTGGATATTAGAAGAAAGCTTCAGAGACTGGAAGG	1200
DB	381	LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly	400
QY	1201	TTACAGAGGCTGAGTTACGTGATCTAGTGAAGAGCGCAGAGAAAGTATATTACAAAGG	1260
DB	401	LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg	420
QY	1261	GAGACAGAGAAGAAAGGGAACAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
DB	421	GluThrGluGluGluLysGluGlnArgLysGluLysGluArgGluGluArgGluGluArg	440
QY	1321	CGTAATAAACCGCAAGAGAGAAGATTTCACCTAAGATCTTGGCTGCAGTGGTTGAAGGAAA	1380
DB	441	ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValGluGlyLys	460
QY	1381	AGCAATACGGAAGAGAGAGAGATTTTAGAAAAATTAGTTCAGGCCCTAGACAGTCAGG	1440
DB	461	SerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly	480
QY	1441	AACCTGGGCAATAGACCCCACTCGACAAGACCAATGTGCATATTGTAAAGAAAGAGA	1500
DB	481	AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly	500
QY	1501	CACCTGGGCAAGAACTGCCCAAGAGGAAACAAAGGACCAAGGATCTTAGCTCTAGAA	1560
DB	501	HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGlu	520
QY	1561	GAAGATAAAGAT 1572	
DB	521	GluAspLysAsp 524	

RESULT 13
 Q8J4V6_9GAMR
 ID Q8J4V6_9GAMR PRELIMINARY; PRT; 1720 AA.
 AC Q8J4V6;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Gag-pol polyprotein.
 OS Porcine endogenous retrovirus B.
 OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
 OC 1-Mammalian type C virus group.
 OX NCBI_TaxID=194959;

[1]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22172947; PubMed=12185278;
 RA Bartosch B., Weiss R.A., Takeuchi Y.;
 RT "PCR-based cloning and immunocytological titration of infectious
 RL porcine endogenous retrovirus subgroup A and B";
 RL J. Gen. Virol. 83:2231-2240(2002).
 DR EMBL; AY099324; AAC29194.1; -; Genomic_DNA.
 DR HSSP; P03355; 116J.
 DR SMR; Q8J4V6; 7-98.
 DR MEROPS; A02.020; -;
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR GO; GO:0019068; P:viral assembly; IEA.
 DR InterPro; IPR003036; Gag_p30.
 DR InterPro; IPR000840; G_retro.matrix.
 DR InterPro; IPR001995; Peptidase_A2_cat.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVTse.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF01140; Gag_MA; 1.
 DR Pfam; PF02093; Gag_p30; 1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT; 1.
 DR Pfam; PF00098; zf_CCHC; 1.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS01075; ASP_PROT_RETROV; 1.
 DR PROSITE; PS50879; RNASE_H; 1.
 DR PROSITE; PS50158; ZF_CCHC; 1.
 KW POLYPEPTIDE.
 SQ SEQUENCE 1720 AA; 193393 MW; 7F9DCC2F5B834D8A CRC64;

Alignment Scores:
 Pred. No.: 1.26e-161 Length: 1720
 Score: 2705.00 Matches: 497
 Percent Similarity: 97.9% Conservative: 16
 Best Local Similarity: 94.8% Mismatches: 11
 Query Match: 95.0% Indels: 0
 DB: 2 Gaps: 0

us-10-723-552-3_COPY_585_2156 (1-1572) x Q8J4V6_9GMR (1-1720)

QY 1 ATGGGACAGCGGTACGACCCCTCTTAGTTGACTCTCGACCACTGGACTGAAGTTAAA 60
 Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20
 QY 61 TCCAGGCTCATATTTGTCAGTTTCAAGTAAGGACCTTGGCAGACTTCTGTGTC 120
 Db 21 SerArgAlaHisAsnLeuSerValGlnValLysGlyProTrpGlnThrPheCysAla 40
 QY 121 TCTGAATGGCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATCTTGAGATT 180
 Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluile 60
 QY 181 ATCTGGCTGTAAAGCAGTTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
 Db 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
 QY 241 CCTTATATCTTACGTGGCAGATTTCGCGAGAGATCTCCGCGCATCGGTTAAACCATGG 300

Db 81 ProTyrlLeuThrTrpGlnAspLeuAlaGluAspProProProProTrpValLysProTrp 100
 QY 301 CTGAATAAGCCAGAAAGCCAGGTCCCGCAATCTCGCTCTGGAGAGAAAAACAAACAC 360
 Db 101 LeuAsnLysProArgLysProGlyProArgileLeuAlaLeuGlyGlyLysAsnLysHis 120
 QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGAGCCACCGCT 420
 Db 121 SerAlaGluLysValGluProSerProArgileTyProGluileGluProProThr 140
 QY 421 TGGCGGNAACCCCAATCTGTTCCTCCCAACCCCTTATCTGGCACAGGTGCCCGAGGGA 480
 Db 141 TrpProGluProGlnProValProProProProTyProAlaGlnGlyAlaValArgly 160
 QY 481 CCTTTTCCCTCTCTGGAGCTCCCGCGGTGAGGACCTGTCTGAGGACCTCGGAGCCGG 540
 Db 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
 QY 541 AGGGCGCCACCCCGGAGCGGACAGACGAGATTCGACATTCACGCTGCGCACGTACGCG 600
 Db 181 ArgGlyAlaThrProGluArgThrAspGluileAlaileLeuProLeuArgThrTyrgly 200
 QY 601 CCTCCACACCGGGGGCCCAATTGACGCCCTCCAGTATTCGCCCTTTCTCTCTGCAGAT 660
 Db 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyTrpProPheSerSerAlaAsp 220
 QY 661 CTCTATAATTGAAAACTAACCATCCCTCTCTCGAGGATCCCAACGCTCACGCGG 720
 Db 221 LeuTyraAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
 QY 721 TTGCTGGAGTCCCTTATGTTCTCTCACCGCTACTTGGGATGATTGTCAACAGCTGCTG 780
 Db 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
 QY 781 CAGACACTCTTCAACCGGAGCGGACGAGAGAAATCTTATAGAGCTAGAAAAATGTT 840
 Db 261 GlnThrLeuPheThrThrGluGluArgGluArgileLeuLeuGluAlaArgLysAsnVal 280
 QY 841 CTGGGGCCGACGGCGGACCCAGCGGTGCAAAATGAGATTGATGATGGGATTTCCCTTA 900
 Db 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluileAspMetGlyPheProLeu 300
 QY 901 ACTCGCCCCGTTGGGACTACAACACCGCTGAAGTAGGGAGAGCTTGAATAATCTATCGC 960
 Db 301 ThrArgProGlyTrpAspTyraAsnThrAlaGluGlyArgGluSerLeuLysileTyArg 320
 QY 961 CAGCTCTGTGGCGGTCTCCGGGCGCTCAAGACGCCCACTAATTTGGCTTAAGTA 1020
 Db 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
 QY 1021 AGAAGAGTGATGAGGAGCCGAATGAACCCCTCTCTGTTTCTTGAGAGCTCTTTGAA 1080
 Db 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
 QY 1081 GCCTTCAGGCGGTACACCCCTTTTGATCCCACTCAGAGGCCCAAAAGCCCTCAGTGCT 1140
 Db 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
 QY 1141 TTGGCTTTTATAGACAGTCTGAGCTTGGATATTAGAAAGAGCTTTCAGAGCTCGAAGG 1200
 Db 381 LeuAlaPheileGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
 QY 1201 TTACAGGAGCTGAGTTTACGTGATCTAGTGAAGGAGGACAGAAAGTATATTACAAAGG 1260
 Db 401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyTrpArgArg 420
 QY 1261 GAGACAGAAGAAGGAGGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 Db 421 GluThrGluGluGluLysGluGlnArgLysGluLysGluArgGluGluArgGluArg 440
 QY 1321 CGTAATAACCGGACAGAGAAATTTGACTTAAGATCTTGGCTGCTGAGTGGTTGAAGGAAA 1380
 Db 441 ArgAspArgArgGlnGluLysAsnLeuThrLysileLeuAlaValValGluGlyLys 460

QY 1381 AGCAATCGGAAGACGAGAGATTTTAGGAAAATTTAGGTTCAGGCCCTAGACAGTCAGGG 1440
Db |||:::|||||
QY 461 SerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
Db |||:::|||||
QY 1441 AACCTGGCAATAGACACCCCTCTGCACAAAGACCAATGTGCATATTTGTAAGAAGAGGA 1500
Db |||:::|||||
QY 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
Db |||:::|||||
QY 1501 CACTGGGCAAGAACTGCCCAAGAGGGAACAAAGACCAAGATCCTAGCTCTAGAA 1560
Db |||:::|||||
QY 501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
Db |||:::|||||
QY 1561 GAAGATAAGAT 1572
Db |||:::|||||
QY 521 GluAspLysAsp 524
Db |||:::|||||
RESULT 14
Q9Q1X4_9GAMR
ID Q9Q1X4_9GAMR PRELIMINARY; PRT: 2378 AA.
AC Q9Q1X4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Type C proviral gag, pol and env genesand LTR (class A, clone
DE 421.
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gamaretrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-type C;
RA Czauderna F., Fischer N., Boller K., Krach U., Kurth R., Toenjes R.R.;
RT "Molecular Characterization of Human-tropic and Replication-competent
RT Porcine Endogenous Retroviruses."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133817; CAB65340.1; -; GenomicDNA.
DR HSRF; P03385; IMOF.
DR SMR; Q9Q1X4; 7-98, 2227-2279.
DR MEROPS; A02.020; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR002050; Env polyproteins.
DR InterPro; IPR008981; FmuLVrecept-bind.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000223; Peptidase_S26A.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR InterPro; IPR001878; ZnF_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF00429; TLV_coat; 1.
DR Pfam; PF00098; zF_CCHC; 1.
DR SMART; SM00343; ZnF_C2HC; 1.
DR PROSITE; PS00141; A5P_PROTEASE; 1.

DR PROSITE; PS0175; ASP PROT RETROV; 1.
DR PROSITE; PS0879; RNASE H; 1.
DR PROSITE; PS00501; SPASE_1_1; UNKNOWN_1.
DR PROSITE; PS0158; 2F_CCHC; 1.
FT CHAIN 1 524 Gag protein.
FT CHAIN 525 1759 Pol protein.
FT CHAIN 1719 2378 Env protein.
SQ SEQUENCE 2378 AA; 267306 MW; 6EC8DAEC1B4ED18A CRC64;

Alignment Scores:
Pred. No.: 1.31e-161 Length: 2378
Score: 2705.00 Matches: 497
Percent Similarity: 97.9% Conservative: 16
Best Local Similarity: 94.8% Mismatches: 11
Query Match: 95.0% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_copy_585_2156 (1-1572) x Q9Q1X4_9GAMR (1-2378)

QY 1 ATGGGACAGCGGTGACGACCCCTCTTAGTTGACTCTCGACCATTTGGACTGAAGTTAAA 60
Db |||:::|||||
QY 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20
Db |||:::|||||
QY 61 TCCAGGGCTCATATTTGTTCAGTTTCAGGTTAAGAGGACCTTGGCAGACTTTCTGTGTC 120
Db |||:::|||||
QY 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAla 40
Db |||:::|||||
QY 121 TCTGAATGGCGGACATTCGATTTGGATGGCATCAGAGGGGACCTTTAAATCTGAGATT 180
Db |||:::|||||
QY 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
Db |||:::|||||
QY 181 ATCTGGCTGTTAAAGCAGTTATTTTTCAGACTCGACCCGGCTCTCATCCCGATCAGGAG 240
Db |||:::|||||
QY 61 IleuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
Db |||:::|||||
QY 241 CCCTATATCTTACGTGGCAGATTGGCAGAGGATCTCCGCCATGGTGTAAACCATGG 300
Db |||:::|||||
QY 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100
Db |||:::|||||
QY 301 CTGAATAGCCAAAGACCGAGTCCCGAATTCGGCTCTTGGAGAGAAAAACAACAC 360
Db |||:::|||||
QY 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
Db |||:::|||||
QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCACCGCT 420
Db |||:::|||||
QY 121 SerAlaGluLysValGluProSerProArgIleIleProGluIleGluGluProProThr 140
Db |||:::|||||
QY 421 TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGGTGCCTCGAGGGGA 480
Db |||:::|||||
QY 141 TrpProGluProGlnProValProProProTyrProAlaGlnGlyAlaValArgGly 160
Db |||:::|||||
QY 481 CCCTTTGCCCTCTCTGGAGCTCCGGCGGTGAGGAGACTGTCTGAGGAGACTCGGAGCGG 540
Db |||:::|||||
QY 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
Db |||:::|||||
QY 541 AGGGGGCCACCCGGGCGGACAGAGATCCGACATTCACCGCTCGCGACGTCACGGC 600
Db |||:::|||||
QY 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
Db |||:::|||||
QY 601 CCTCCACACCGGGGGCCAAATTGTCAGCCCTCCAGATTTGGCCCTTTTCTTCGACAT 660
Db |||:::|||||
QY 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220
Db |||:::|||||
QY 661 CTCTATATTTGGAAAACTAACCATCCCTCTTCGGAGGATCCCAACGCTCAGCGGG 720
Db |||:::|||||
QY 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
Db |||:::|||||
QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACCAGCCTACTCTGGGATGATTGTCAACAGCTGCTG 780
Db |||:::|||||
QY 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
Db |||:::|||||
QY 781 CAGACACTCTTTCAACAACCGGAGGAGAGAAATTTATTAGAGGCTAGAAAAAATGTT 840
Db |||:::|||||


```
Db      201  ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaasp 220
Qy      661  CTCATATAATTGGAATAACTAAACATCCCTCTTCGAGAGATCCCCAAACGCTCACGGGG 720
Db      221  LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
Qy      721  TTGTTGGAGTCCCTTATCTCTCTCACAGCTACTTGGGATGATTGTCAACAGCTGCTG 780
Db      241  LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
Qy      781  CAGACACTCTTTCACACCGAGAGCGAGAGAGAAATTTCTATTAGAGGCTAGAAAAAATGTT 840
Db      261  GlnThrLeuPheThrThrGluGluArgGluArgGluLeuLeuGluAlaLysLysAsnVal 280
Qy      841  CCTGGGCGGAGCGGCGGACCCACGCGGTTCGAAATGAGATTGACATGGGATTTCCCTTA 900
Db      281  ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
Qy      901  ACTGCCCGCGTTCGGACTACAAACAGCTCAAGCTAGGAGAGCTTGAAAAATCTATCGC 960
Db      301  ThrArgProGlyTrpAspTrpAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
Qy      961  CAGGCTCTGTGGGGGTCTCCGGGGCGCTCAAGACGGCCCTAAATTTGGCTAAGGTA 1020
Db      321  GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
Qy      1021  AGAGAGTGATGCGAGGACCGAATGAACCCCTCTGTTTTCTTTGAGAGGCTCTTGAA 1080
Db      341  ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
Qy      1081  GCCTTCAGCGGTACACCCCTTTTGATCCACCTCAGAGGCCCAAAAGCCCTCAGTGGCT 1140
Db      361  AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
Qy      1141  TTGGCTTTATAGACAGCTCAGCTTGATGATATTAGAAAGAGCTTCAGAGACTGGAAGG 1200
Db      381  LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
Qy      1201  TTACAGGAGCTGAGTTACGTGATCTACTGAGAGGCGAGAGAAAGTATATATTACAAAGG 1260
Db      401  LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrArgArg 420
Qy      1261  GAGACAGAGAGAAAGGAAACAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db      421  GluThrGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 440
Qy      1321  CGTAATAACCGCAAGAGAGAAATTTGACTAAGATCTTGGCTGCGAGTGGTTGAAAGGAAA 1380
Db      441  ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValGluGlyLys 460
Qy      1381  AGCAATACGGAAGAGAGAGAGATTTTAGGAAAATTAGGTCAGGCCCTAGACAGTCAGGG 1440
Db      461  SerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
Qy      1441  AACCTGGGCAATAGGACCCCACTCGACAAGGACCAATGCTGATATTGTAAAGAAAGAGGA 1500
Db      481  AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
Qy      1501  CACTGGGCAAGGAATGCCCCAAGAGGGAACCAAGGACCAAGGATCTAGCTCTAGAA 1560
Db      501  HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
Qy      1561  GAAGATAAGAT 1572
Db      521  GluAspLysAsp 524
```

Search completed: February 14, 2006, 14:45:47
Job time : 372.026 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - protein search, using frame_plus_n2p model
Run on: February 14, 2006, 12:52:14 ; Search time 41.8246 Seconds
(without alignments)
3302.856 Million cell updates/sec
Title: US-10-723-552-3_COPY_585_2156
Perfect score: 2848
Sequence: 1 ATGGGACAGCGGTGACGAC.....CTCTAGAGAGATAAAGAT 1572

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10723552/runat 14022006 125139 12833/app query.fasta_1
-DB=A Geneseq -QFMT=fastan -SUPFIX=rag -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10723552 @CGN 1 1 734 @runat 14022006 125139 12833 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	2817	98.9	524	2	AAW32096	Aw32096 Miniature
2	2817	98.9	524	4	AAB73285	Aab73285 Retroviral
3	2817	98.9	524	5	AAO19011	Aao19011 Porcine e
4	2817	98.9	524	8	ADS73446	Ads73446 Swine ret
5	2813	98.8	744	8	ADS73450	Ads73450 Swine ret
6	2705	95.0	524	5	AAO19009	Aao19009 Porcine e
7	2705	95.0	524	5	AAO19010	Aao19010 Porcine e
8	2702	94.9	524	2	AAW32091	Aw32091 Porcine r
9	2702	94.9	524	4	AAB73282	Aab73282 Defective

10	2702	94.9	524	8	ADS73441	AdS73441	Swine ret
11	2640	92.7	524	2	AAW39271	Porcine r	
12	2640	92.7	524	4	AAW70629	Porcine e	
13	1733	60.8	522	2	AAW81571	Mus dunni	
14	1733	60.8	622	2	AAW81570	Mus dunni	
15	1380.5	48.5	538	3	AAAB10043	MMLV gag	
16	1380.5	48.5	538	7	ADH76469	Murine le	
17	1380.5	48.5	648	2	AAW17946	MMLV gag	
18	1380.5	48.5	1737	3	AAAB10044	MMLV gag	
19	1379.5	48.4	1784	2	AAAR05898	Gene prod	
20	1377.5	48.4	1784	2	AAAR94427	FelV F6A	
21	1374	48.2	580	2	AAW68402	FelV-A ga	
22	1371	48.1	537	4	AAAB49467	Canine re	
23	1371	48.1	625	4	AAAB49466	Canine re	
24	1359	47.7	508	8	ADDO3718	Feline le	
25	1357	47.6	537	2	AAAR75188	Osteoindu	
26	1348.5	47.3	509	8	ADDO3720	Feline le	
27	930.5	32.7	499	5	ABW79874	Spleen ne	
28	842	29.6	441	7	ADH77072	Friend mu	
29	842	29.6	529	7	ADH77071	Friend mu	
30	832.5	29.2	253	4	AAAM50120	Feline le	
31	826	29.0	263	2	AAAR80533	Moloney m	
32	824	28.9	263	2	AAAR80537	Moloney m	
33	820	28.8	263	2	AAAR80535	Moloney m	
34	668	23.5	523	8	ADW50749	Human pro	
35	553	19.4	405	4	ABG07968	Novel hum	
36	553	19.4	430	4	ABG08938	Novel hum	
37	553	19.4	1577	4	ABG06475	Novel hum	
38	553	19.4	1577	4	ABG27123	Novel hum	
39	553	19.4	1650	4	AAAM39205	Human pol	
40	543.5	19.1	611	4	AAAM50122	Feline le	
41	542.5	19.0	878	4	ABG06484	Novel hum	
42	526.5	18.5	688	4	ABG08930	Novel hum	
43	510.5	17.9	205	3	ABAB39242	Human sec	
44	478	16.8	505	4	ABG15313	Novel hum	
45	475.5	16.7	208	8	ABOS9520	Human gen	

ALIGNMENTS

RESULT 1
AAW32096
ID AAW32096 standard; protein; 524 AA.
XX
AC AAW32096;
XX
DT 27-AUG-2003 (revised)
DT 09-FEB-1998 (first entry)
XX
DE Miniature swine retrovirus GAG protein.
XX
KW Retrovirus; porcine; GAG protein; xenotransplantation; infectious;
KW provirus; organ transplant; donor; activated virus; PCR.
XX
OS Pig endogenous retrovirus.
XX
FH Key Location/Qualifiers
FT Protein 1..524
FT /label= GAG_protein
XX
PN WO9721836-A1.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96WO-US019680.
XX
PR 14-DEC-1995; 95US-00572645.
XX
PA (GEO) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI; 1997-332804/30.

DR N-PSDB; AAT74884.
 XX New nucleic acid from porcine retro:viruses - used for detecting viruses
 PT in transplant or other tissue and for assessing risk of transmitting
 PT infection to graft recipient.
 XX
 PS Claim 22; Fig 3; 128pp; English.

XX This is a porcine retrovirus from miniature swine containing the coding
 CC region for a putative viral GAG protein. This sequence and PCR fragments
 CC generated from the sequence (see AAT74812-T74882) could be used to screen
 CC organs for porcine retroviruses prior to xenotransplantation.
 CC Transplantation can increase the likelihood of retroviral activation if
 CC intact and infectious proviruses are present. The porcine retroviral
 CC sequence can be used to generate probes to determine the level (e.g. copy
 CC number) of intact (i.e. potentially replicating) porcine provirus
 CC sequences in a strain of xenograft transplantation donors. It can be used
 CC to detect mutations, genetic lesions or viral recombinants and also to
 CC determine the histological localisation of activated retrovirus. Using
 CC Polymerase Chain Reaction DNA Quantitation (PDQ) on blood mononuclear
 CC cells, infectivity titration and susceptibility testing can be performed.
 CC Ultimately animal donors without intact porcine retroviral sequences or a
 CC lower copy number of viral elements could be selected. (Updated on 27-AUG
 CC -2003 to correct OS field.)

SQ Sequence 524 AA;

Alignment Scores:

Pred. No.: 4,09e-241 Length: 524
 Score: 2817.00 Matches: 524
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 98.9% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x AAW32096 (1-524)

QY 1 ATGGGACAGAGGTCAGACCCCTTCTAGTTTGACTCTCGACCATTCGAGTTAA 60
 DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluVallys 20
 QY 61 TCCAGGGCTCATAATTTCTCAGTTTCAAGTAAAGGACCTTGGCAGACTTTCTGTGC 120
 DB 21 SerArgAlaHisAsnLeuSerValGlnVallylsysGlyProTrpGlnThrPheCysVal 40
 QY 121 TCTGAATGCCGACATTCGATGTGGATGGCCATCAGAGGGGACCTTTAATTCGAGATT 180
 DB 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluile 60
 QY 181 ATCTGGCTGTAAAGCAGTTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
 DB 61 IleLeuAlaValLysAlaValIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
 QY 241 CCCTATATCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATGGGTTAAACCATGG 300
 DB 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProProTrpValysProTrp 100
 QY 301 CTGAATAAGCCAAGAAACCGAGTCCCGAATTCGGCTCTTGGAGAGAAACAAACAC 360
 DB 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
 QY 361 TCGGCTGAAAAGTCAGCCCTCTCTCATATCTACCCCGAGATTGAGGCCACCGGCT 420
 DB 121 SerAlaGluLysValLysProSerProHisIleTrpProGluIleGluProProAla 140
 QY 421 TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGGTGGCCGAGGGGA 480
 DB 141 TrpProGluProGlnSerValProProProTrpTyrLeuAlaGlnGlnIleAlaalaArgGly 160
 QY 481 CCCTTTGCCCTCTCGGAGCTCCGGCGTGGAGGACCTGTGCGAGGACCTCGGAGCCGG 540
 DB 161 ProPheAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180

RESULT 2

QY 541 AGGGGCGCCACCCCGAGCGGACAGACAGATCGACATTACCGCTGCGCACGTCACGGC 600
 DB 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly 200
 QY 601 CCTCCACACCGGGGCGCAATTCGAGCCCTCCAGTATTGGCCCTTTTCTTCGACAGAT 660
 DB 201 ProProThrProGlyGlnLeuGlnProLeuGlnTrpTrpProPheSerSerAlaAsp 220
 QY 661 CTCTATAATTGGAAAACTAACCATCCCTTCTCGGAGGATCCCAACGCTCACCAGG 720
 DB 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
 QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACAGCCTACTTGGGATGATTGTCAACAGCTGCTG 780
 DB 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
 QY 781 CAGACACTCTTTCACACCGAGGACGAGAGAGAAATCTTATTAGAGGCTAGAAAAATGTT 840
 DB 261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
 QY 841 CCTGGGCGGACGCGGCGACCCACCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTA 900
 DB 281 ProGlyAlaAspGlyArgProThrArgLeuGlnAsnGluIleAspMetGlyPheProLeu 300
 QY 901 ACTGCGCCCGGTTGGGACTACACACGCTCAAGCTAGGAGAGCTTGAAAATCTATCGC 960
 DB 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
 QY 961 CAGGCTCTGTGGCGGGTCTCCGGGCGCCTCAAGACGGCCCACTAATTTGGCTAAGGTA 1020
 DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
 QY 1021 AGAGAAGTGATGCAGGACCGAATGAACCCCTCTGTTTTCTTGAGGCTCTTCGAA 1080
 DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuGlu 360
 QY 1081 GCCTTCAGGCGGTACACCCCTTTTGTATCCACCTCAGAGGCCCAAAAGCCTCAGTGGCT 1140
 DB 361 AlaPheArgArgTyrThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
 QY 1141 TTGCGCTTATPAGACAGTCAAGCTTGGATATTAGAAGAAAGCTTCAGAGACTGGAAGG 1200
 DB 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
 QY 1201 TTACAGAGGCTGAGTTACGTGATCTAGTGAGGCGCAGAGAAAGTATATACAAAAGG 1260
 DB 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420
 QY 1261 GAGACAGAAGAAAGGAAACAAAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 1320
 DB 421 GluThrGluGluGluArgGluGluGluGluGluGluGluGluGluGluGluGluArg 440
 QY 1321 CGTATAAACCGCAGAGAGAAATTTGACTAAGATCTTGGCTGCGAGTGGTTGAAAGGAAA 1380
 DB 441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaValValGluGlyLys 460
 QY 1381 AGCAATACGGAACAGAGAGAGATTTTAGGAAAAATAGGTCAGGCCCTTAGCAGTCAGG 1440
 DB 461 SerAsnThrGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
 QY 1441 AACCTGGGCAATAGGACCCCTCAGCAAGACCAATGTGCATATTGTAAAGAAAGAGA 1500
 DB 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGly 500
 QY 1501 CACTGGGCAAGGAACTGCCCAAGAGAGGAAACAAAGGACCAAGGATCCTAGCTCTAGAA 1560
 DB 501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProArgIleLeuAlaLeuGlu 520
 QY 1561 GAAGATAAGAT 1572
 DB 521 GluAspLysAsp 524

AAB73285

ID AAB73285 standard; protein; 524 AA.

AC AAB73285;

XX 23-MAY-2001 (first entry)

XX Retroviral protein #1 found in miniature swine.

XX Retrovirus; graft transplantation; xenotransplantation; miniature swine.

XX Unidentified.

XX US6190861-B1.

XX 20-FEB-2001.

XX 13-DEC-1996; 96US-00766528.

XX 14-DEC-1995; 95US-00572645.

XX (GEO) GEN HOSPITAL CORP.

XX Fishman JA;

XX WPI; 2001-256211/26.

XX N-PSDB; AAF77727.

XX Assessing risk of endogenous retroviruses in clinical practice and in
XX xenotransplantation, comprises using probe sequences derived from swine
XX or miniature swine retroviral genome.

XX Disclosure; Fig 3; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue
XX for the presence or expression of a retrovirus (RV), comprising
XX contacting a target nucleic acid from the cell or tissue with a second
XX nucleic acid from the present invention (e.g. AAF7727 or a fragment
XX thereof). The method is useful for RV detection and to assess graft
XX transplantation risk. Screening of animals allows the elimination of
XX donors with active replication of known viruses. Inactive proviruses can
XX be detected and inactivated, allowing identification and elimination of
XX potential human pathogens derived from swine in a manner not possible in
XX the outbred human organ donor population and is important to the
XX development of human xenotransplantation

XX Sequence 524 AA;

Alignment Scores:

Pred. No.: 4.09e-241 Length: 524
Score: 2817.00 Matches: 524
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.9% Indels: 0
DB: 4 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x AAB73285 (1-524)

QY 1 ATGGGACAGCGGTGACGACCCCTTGTAGTTGACTCTCGACCATTCGACTGAAGTTAAA 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluVallys 20
QY 61 TCCAGGGCTCAATAATTTGTTCAGTTTCAGTTAAGAGGACCTTCGACAGACTTCTGTGTC 120
DB 21 SerArgalaHisAsnLeuSerValGlnVallysGlyProTrpGlnThrPheCysVal 40
QY 121 TCTGAATGGCCGACATTTCGATGTTGGATGCGCATCAGAGGGGACCTTTAATTCGAGATT 180
DB 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluile 60
QY 181 ATCCCTGGCTGTTAAGCAGTTATTTTCAGACTGACCCGGCTCTCATCCGATCAGAG 240
DB 61 IleLeuAlaVallysAlaValIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80

QY 241 CCCTATATCTTACGTGGCAAGATTTCGACAGGATCTCTCGCCATGGGTTAAACCATGG 300
DB 81 ProTyrlleLeuThrTrpGlnAspLeuAlaGluAspProProTrpVallysProTrp 100
QY 301 CTGAATAAGCCCAAGAACCCAGGTCCCGAATTCCTGGCTCTTGGAGAGAAAAACACAC 360
DB 101 LeuAsnLysProArgLysProGlyProArgileLeuAlaLeuGlyGluLysAsnLysHis 120
QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCATATCTTACCCCGAGATTTCGAGAGCCACCGGCT 420
DB 121 SerAlaGluLysVallysProSerProHisileTyProGluileGluGluProProAla 140
QY 421 TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGGTCCCGCAGGGGA 480
DB 141 TrpProGluProGlnSerValProProProProTyrlleuAlaGlnGlyAlaAlaArgGly 160
QY 481 CCTTTGGCCCTCTCTGGAGCTCCCGCGGTGGAGGACCTCTGTCAGGGACTCGGAGCCGG 540
DB 161 ProPheAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
QY 541 AGGGGCGCCACCCCGAGCGGACAGACGAGATCCGACATTACCGCTCGCGACCTACCGC 600
DB 181 ArgGlyAlaThrProGluArgThrAspGluileAlaThrLeuProLeuArgThrTyrlly 200
QY 601 CCTCCACACCGGGGGGCCAATTTCAGCCCTTCCAGTATTCGCCCTTTTCTTCTGCAGAT 660
DB 201 ProProThrProGlyGlyGlnLeuGlnProLeuGlnTyTrpProPheSerSerAlaAsp 220
QY 661 CTCTATATTCGAAAAAATAACCATCCCTCTTTCGGAGGATCCCAACCGCTCACGGGG 720
DB 221 LeuTyraAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
QY 721 TTGGTGAGTCCCTTATGTTCTCTCACAGCCTACTTGGGATGATTGTCAACAGCTGCTG 780
DB 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
QY 781 CAGACACTTCTCAACCCGAGGCGAGAGAGAAATTCATTAGAGGCTAGAAAAAATGTT 840
DB 261 GlnThrLeuPheThrThrGluGluArgGluArgileLeuLeuGluAlaArgLysAsnVal 280
QY 841 CTGGGGCCGACGGGCCACCCAGCGGTTGCAAAATGAGATTGACATGGGATTCCTCTTA 900
DB 281 ProGlyAlaAspGlyArgProThrArgLeuGlnAsnGluileAspMetGlyPheProLeu 300
QY 901 ACTCGCCCGGTTGGGACTACACAGCTGAGGTAGGAGAGCTTGAATAATCTATCGC 960
DB 301 ThrArgProGlyTrpAspTyraAsnThrAlaGluGlyArgGluSerLeuLysileTyArg 320
QY 961 CAGGCTCTGTTGGCGGCTCTCCGGGGCGCTCAAGAGCGGCCCACTAATTTGGCTAAGGTA 1020
DB 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
QY 1021 AGAGAGTGTATGACGGGACCCGAATGAACCCCTCTCTTTTCTTTGAGAGGCTCTTGAA 1080
DB 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuGlu 360
QY 1081 GCCTTCAGGGGTACACCCCTTTTGTATCCCACTCAGAGGCCCAAAAGCTCAGTGGCT 1140
DB 361 AlaPheArgArgTyTrpProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
QY 1141 TTGGCTCTTATAGACAGCTCAGCCCTTCGATATTAGAAGAAGCTTCAGAGACTCGAAGG 1200
DB 381 LeuAlaPheileGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
QY 1201 TTAAGAGGCTGAGTTTACCTGATCTAGTCAAGAGGACAGAGAAAGTATATTACAAAGG 1260
DB 401 LeuGlnGluAlaGluLeuArgAspLeuVallysGluAlaGluLysValTyTrpLysArg 420
QY 1261 GAGACAGACAGAAAGGGGACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluGluArgGluGluArg 440

QY 1321 CGTATAAACGCGAAGAGAAATTGACTTAAGATCTTGGCTGCGAGTGGTTGAAGGGAAA 1380
Db |||||||
QY 441 ArgAsnLysAsgGlnGluYsAsnLeuThrLysIleLeuAlaValAluGluGlyLys 460
Db |||||||
QY 1381 AGCAATACGGAAGAGAGAGATTATTTAGGAAAATTTAGGTGAGGCCCTTAGACAGTCAGGG 1440
Db |||||||
QY 461 SerAsnThrGluAsgGluAsgPheArgLysIleAsgSerGlyProAsgGlnSerGly 480
QY 1441 AACCTGGGCAATAGGACCCCACTGCAAGACCAATGTGCATATTGTAAGAAGAGGGA 1500
Db |||||||
QY 481 AsnLeuGlyAsnAsgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluAsgGly 500
QY 1501 CACTGGGCAAGCACTGCCCAAGAGGGAACCAAGCAAGGATCTAGCTAGAA 1560
Db |||||||
QY 501 HisTrpAlaAsgAsnCysProLysGlyAsnLysGlyProAsgLysGlnLeuAlaLeuGlu 520
Db |||||||
QY 1561 GAAGATAAAGAT 1572
Db |||||||
QY 521 GluAsgLysAsg 524
Db |||||||

RESULT 3

AAO19011
ID AAO19011 standard; protein; 524 AA.
XX
AC AAO19011;
XX
DT 07-NOV-2002 (first entry)
XX
DE Porcine endogenous retrovirus type C gag.
XX
KW PERV; protease; virucide; infection; xenotransplant; gag.
XX
OS Porcine endogenous retrovirus.
XX
PN DE10138528-A1.
XX
PD 04-JUL-2002.
XX
PF 06-AUG-2001; 2001DE-01038528.
XX
PR 23-DEC-2000; 2000DE-01064897.
XX
PA (UYMU-) UNIV MUENCHEN MAXIMILIANS LUDWIG.
PI Von Der Helm K, Blusch JH, Seelmeir S;
XX
DR WPI; 2002-609974/66.
XX
DR N-PSDB; AAL49335.
XX
PT New nucleic acid, useful for identifying specific inhibitors for
PT controlling infection following xenotransplantation, encodes porcine
PT endogenous retrovirus protease.
XX
PS Disclosure; Page 18-19; 30pp; German.
XX
CC The present invention relates to proteases and their coding sequences
CC from porcine endogenous retroviruses (PERVs) type A, B and C. The coding
CC sequences are useful in screening for inhibitors of proteases,
CC particularly where retroviral. The inhibitors are potentially useful for
CC treating or preventing PERV infections in humans, specifically those who
CC have received a pig organ transplant. The present sequence is a
CC retroviral gag protein
XX
SQ Sequence 524 AA;

Alignment Scores:

Pred. No.: 4,098-241 Length: 524
Scores: 2817.00 Matches: 524
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.9% Indels: 0
DB: 5 Gaps: 0

US-10-723-552-3_copy_585_2156 (1-1572) x AAO19011 (1-524)
QY 1 ATGGGACAGACGGTGCAGACCCCTCTTAGTTTGACTCTCGACCATTTGGACTGAAGTTAAA 60
Db |||||||
QY 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAsePHisTrpThrGluValLys 20
QY 61 TCCAGGCGTCATAATTTGTTCAGTTTCAGTTTAAAGAGGACCTTGGCAGACTTCTGTGTC 120
Db |||||||
QY 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 40
QY 121 TCTGAATGGCGCAGCATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAAATTCAGATT 180
Db |||||||
QY 41 SerGlnTrpProThrPheAsePValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
QY 181 ATCTGCGCTGTTAAAGCAGATTATTTTCAGACTGACCCGGCTCTCATCCGATCAGGAG 240
Db |||||||
QY 61 IleLeuAlaValLysAlaValIlePheGlnThrGlyProGlySerHisProAsePGLnGlu 80
QY 241 CCTTATATCTTACGTGCAAGATTTCGACAGAGATCTCCGCCATGGGTTAAACCATGG 300
Db |||||||
QY 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAsePProProProTrpValLysProTrp 100
QY 301 CTGAATAAGCAAGAAACCCAGGTCCCGAATTCGGCTCTTGAGAGAAAAACAAACAC 360
Db |||||||
QY 101 LeuAsnLysProAsgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCCTCATATCTACCCCGAGATTGAGGACCCCGCT 420
Db |||||||
QY 121 SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProProAla 140
QY 421 TGGCGGGAACCCCAATCTGTTCCCAACCCCTTATCTGGCACAGGGTCCCGCAGGGGA 480
Db |||||||
QY 141 TrpProGluProGlnSerValProProProTyrLeuAlaGlnGlyAlaAlaAsgGly 160
QY 481 CCCTTTGCCCTCCTGGAGCTCCGGCGGTGGAGGACCTGTCTGAGGAGACTCGAGCGG 540
Db |||||||
QY 161 PropheAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg 180
QY 541 AGGGCGGCCACCCCGGAGCGGACAGACAGATCCGACATTACCGCTGCGCACGTACGGC 600
Db |||||||
QY 181 ArgGlyAlaThrProGluAsgThrAsePGLuIleAlaThrLeuProLeuAsgThrTyrGly 200
QY 601 CCTCCACACCGGGGGCCAAATTCAGCCCTCCAGTATTGGCCCTTTCTTCTGCAGAT 660
Db |||||||
QY 201 ProProThrProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAseP 220
QY 661 CTCTATAATTGGAAAACTAACCATCCCTTTCTGGAGGATCCCAACGCTCACGGGG 720
Db |||||||
QY 221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAsePProGlnArgLeuThrGly 240
QY 721 TTGGTGGAGTCCCTTATGTTCTCTCACCAGCTACTTGGGATGATTGTCAACAGCTGCTG 780
Db |||||||
QY 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAsePAspCysGlnGlnLeuLeu 260
QY 781 CAGACACTCTTCAACAACCGAGGACGAGAGAATTTCTATTAGAGGCTAGAAAAAATGTT 840
Db |||||||
QY 261 GlnThrLeuPheThrThrGluGluAsgGluAsgLeuLeuGluAlaAsgLysAsnVal 280
QY 841 CCTGGGCGGACGGGCGGACCCACCGGTGCAAAATGAGATTGACATGGGATTCCCTTA 900
Db |||||||
QY 281 ProGlyAlaAsePGLysArgProThrArgLeuGlnAsnGluIleAsePMetGlyPheProLeu 300
QY 901 ACTGCCCCGGTTGGGACTACAACCGCTGAAGGTAGGAGAGCTTGAATCTATCGC 960
Db |||||||
QY 301 ThrArgProGlyTrpAsePTrpAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
QY 961 CAGGCTCTGGTGGCGGGTCTCCGGCGGCTCAAGACGGCCCACTAAATTTGGCTAAGGTA 1020
Db |||||||
QY 321 GlnAlaLeuValAlaGlyLeuAsgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
QY 1021 AGAGAAGTGATGACGAGGACCGAATGAACCCCTCTCTGTTTTCTTGAGAGCGCTCTGAA 1080
Db |||||||
QY 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluAsgLeuLeuGlu 360

QY 1081 GCCTTCAGCGGTACACCCCTTTTGCATCCCTCAGAGGCCCAAAAAGCCTCAGTGCT 1140
Db |||||
QY 361 AlaPheArgTyThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
Db |||||
QY 1141 TTGGCCTTTATAGGACAGTCAGCCTTGGATATAGAAAGAGCTTCAGAGCTGAAGG 1200
Db |||||
QY 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLeuGlnArgLeuGly 400
Db |||||
QY 1201 TTACAGGAGGCTGAGTACGTGATCTAGTGAAGGAGGACAGAAAAGTATATTACAAAAGG 1260
Db |||||
QY 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420
Db |||||
QY 1261 GAGACAGAGAAGAAAGGAAACAAAGAAAGAGAGAGAAAGAGGAAAGGAGGAAAGA 1320
Db |||||
QY 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluArgGluArgGluArg 440
QY 1321 CGTATATACGCGACAGAGAGAAATTTGACTAAGATCTTGGCTGCAGTGTTCAGAGGAAA 1380
Db |||||
QY 441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValGluGlyLys 460
QY 1381 AGCAATACGGAAGAGAGAGAGATTTTAGGAAAATTAGGTCAAGGCCCTAGACAGTCAGG 1440
Db |||||
QY 461 SerAsnThrGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
QY 1441 AACCTGGCAATAGACCCCACTCGACAAGGACCAATGTGCATATTGTAAAGAAAGAGGA 1500
Db |||||
QY 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGly 500
QY 1501 CACTGGGCAAGAACTGCCCAAGAGGAAACAAAGACCAAGGATCCTAGCTTAGAA 1560
Db |||||
QY 501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProArgIleLeuAlaLeuGlu 520
QY 1561 GAAGATAAAGAT 1572
Db |||||
QY 521 GluAspLysAsp 524

RESULT 4

ID ADS73446
AD573446 standard; protein; 524 AA.

XX AC ADS73446;

XX DT 16-DEC-2004 (first entry)

XX DE Swine retroviral gag protein.

XX KW Swine retroviral protein; immunosuppressive; gene therapy; gag protein.

XX OS Porcine endogenous retrovirus.

XX PN US2004185435-A1.

XX PD 23-SEP-2004.

XX PF 26-NOV-2003; 2003US-00723552.

XX PR 14-DEC-1995; 95US-00572645.

XX PR 13-DEC-1996; 96US-00766528.

XX PR 14-SEP-2000; 2000US-00661858.

XX FA (GEHO) GEN HOSPITAL CORP.

XX XX Fishman JA;

XX DR WPI; 2004-689179/67.

XX DR N-ESDB; ADS73369.

XX PT New porcine retroviral polypeptide encoded by a nucleic acid, useful in

XX PT evaluating an immunosuppressive treatment for the ability to activate a

XX XX retrovirus, such as an endogenous porcine retrovirus.

XX PS Claim 8; Fig 3; 83pp; English.

XX The present invention relates to the swine retroviral polypeptides and
CC their encoding nucleic acids. The methods and compositions of the present
CC invention are useful for screening a cell or tissue, e.g. a heart, lung,
CC liver, bone marrow, kidney, brain cells, neural tissue, pancreas and
CC intestinal tissue xenograft, for the presence or expression of a swine or
CC miniature swine retrovirus or retroviral sequence. The invention is also
CC useful in evaluating an immunosuppressive treatment for the ability to
CC activate a retrovirus such as an endogenous porcine retrovirus. The
CC invention is also useful in gene therapy. The present sequence is the
CC swine retroviral gag protein.
XX

SQ Sequence 524 AA;

Alignment Scores:

Pred. No.:	4.09e-241	Length:	524
Score:	2817.00	Matches:	524
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	98.9%	Indels:	0
DB:	8	Gaps:	0

US-10-723-552-3_COPY_585_2156 (1-1572) x ADS73446 (1-524)

QY	1	ATGGGACAGACGGTGACGACCCCTCTTAGTTTGA	CTCTCGACCAATGGACTGAAGTTAAA	60
Db	1	MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValLys	20	
QY	61	TCGAGGCTCATAATTTGTCTCAGTTACAGTTAAGAGGACCTTGGCAGACTTCTGTGTC	120	
Db	21	SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal	40	
QY	121	TCTCAATGGCGCACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATCTCAGATT	180	
Db	41	SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle	60	
QY	181	ATCCTGGCTGTAAAGCAGTTATTTTTCAGACTCGACCCGCTCTCATCCGATCAGGAG	240	
Db	61	IleLeuAlaValLysAlaValIlePheGlnThrGlyProGlySerHisProAspGlnGlu	80	
QY	241	CCCTATATCCTTAGCTGGCAAGATTTCGACAGGATCCTCGCCATGGTTAAACCATGG	300	
Db	81	ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp	100	
QY	301	CTGAATAAGCCAAAGACCCAGGTCGCCGAATTCGTCTCTTGAGAGAGAAAAACAACAC	360	
Db	101	LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGlyLysAsnLysHis	120	
QY	361	TCGGCTCAAAAAGTCAAGCCCTCTCTCATATATCTACCCCGAGATTGAGGACCCCGCT	420	
Db	121	SerAlaGluLysValLysProSerProHisIleTyrProGluIleGluGluProAla	140	
QY	421	TGGCCGGAACCCCAATCTGTTCCCCACCCCTTATCTGGCAGAGGTGCGCGAGGGA	480	
Db	141	TrpProGluProGlnSerValProProProTyrLeuAlaGlnGlyAlaAlaArgGly	160	
QY	481	CCCTTTGCCCTCTCTCGAGCTCCGGCGGTGGAGGACCTGCTGCAGGACTCGAGCGCG	540	
Db	161	ProPheAlaProProGlyAlaProAlaValGluGlyProAlaAlaGlyThrArgSerArg	180	
QY	541	AGGGCGCCACCCCGGAGCGGACAGACGAGATCGCAGATTACCGCTGCGCACCTACGGC	600	
Db	181	ArgGlyAlaThrProGluArgThrAspGluIleAlaThrLeuProLeuArgThrTyrGly	200	
QY	601	CTTCCGACACCGGGGGCCAAATTGACGCCCTTCCAGTATTGGCCCTTTTCTTTCAGAT	660	
Db	201	ProProThrProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp	220	
QY	661	CTCTATAATTTGGAAACTAACCATCCCTTTCTCGAGGATCCCAACCCCTCACCGGG	720	
Db	221	LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly	240	
QY	721	TTGGTGGAGTCCCTTATGTTCTCTCACGACCTACTTGGGATGATTGTCAACAGCTGCTG	780	

```
|||||
241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
QY 781 CAGACACTTTCACAAACGAGAGCGAGAGAAATTCATTAGAGCGCTAGAAAAATGTT 840
Db 261 GlnThrLeuPheThrThrGluGluArgGluArgGluLeuLeuGluAlaArgLysAsnVal 280
QY 841 CCTGGGCGGAGCGGCGGACCCACGCGTTGCMAATGAGATTGACATGGGATTCCTTA 900
Db 281 ProGlyAlaAspGlyArgProThrArgLeuGlnAsnGluIleAspMetGlyPheProLeu 300
QY 901 ACTGCCCCCGGTGGCAGCTACAAACCGCTGAAGCTAGCGAGCTTGAAATCTATCGC 960
Db 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
QY 961 CAGGCTCTGGTGGCGGCTCTCGGGCGCCTCAAGACGCGCCACTAAATTTGGCTAAGTA 1020
Db 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
QY 1021 AGAAGAGTGTGCAGGAGCGGACCGAATCAACCCCTCTGTTTCTTGAGAGGCTCTTGAA 1080
Db 341 ArgGluValMetGlnGlyProAsnGluProSerValPheLeuGluArgLeuLeuGlu 360
QY 1081 GCCTTCAGCGGTACACCCCTTTTGATCCACCTCAGAGGCCCAAAAGCCTCAGTGCT 1140
Db 361 AlaPheArgArgTyrThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
QY 1141 TTGGCCTTTATAGACAGTCAGCCTTGATATTAGAAAGAGCTTCAGAGCTGGAAGG 1200
Db 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGly 400
QY 1201 TTACAGGAGGCTGAGTTACGTGATCTAGTAGGAGCGCAGAAAGTATATTACAAAGG 1260
Db 401 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 420
QY 1261 GAGCAGAGAAGAAAGGAAACAAAGAAAGAGAGAGAAAGAGAGAAAGGAGGAGAAAGA 1320
Db 421 GluThrGluGluGluArgGluGlnArgLysGluArgGluArgGluGluArgGluArg 440
QY 1321 CGTAATAACCGCAAGAGAAATTTGACTTAAGATCTTTGGCTGCGAGTGGTGAAGGAAA 1380
Db 441 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLys 460
QY 1381 ACCATACGGAAGAGAGAGATTTTAGGAAATTAGCTCAGCGCCCTAGACAGTCAGGG 1440
Db 461 SerAsnThrGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480
QY 1441 AACCTGGGCAATAGCAGCCCACTCCAGCAGGACCAATGTCATATTGTAAAGAAAGAGA 1500
Db 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGly 500
QY 1501 CACTCGGCAAGGAATGCCCCCAAGAGGAAACAAAGGACCAAGGATCTTAGCTTAGAA 1560
Db 501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProArgIleLeuAlaLeuGlu 520
QY 1561 GAAGATAAAGAT 1572
Db 521 GluAspLysAsp 524
RESULT 5
ID ADS73450 standard; protein; 744 AA..
XX AC ADS73450;
XX DT 16-DEC-2004 (first entry)
XX DE Swine retroviral (Tsukuba-1) gag protein.
XX KW Swine retroviral protein; immunosuppressive; gene therapy; gag protein.
XX OS Porcine endogenous retrovirus.
XX
```

```
Key Location/Qualifiers
FH Misc-difference 53
FT /note= "Encoded by TGA"
FT Misc-difference 134
FT /note= "Encoded by TGA"
XX US2004185435-A1.
PN 23-SEP-2004.
XX 26-NOV-2003; 2003US-00723552.
XX 14-DEC-1995; 95US-00572645.
XX 13-DEC-1996; 96US-00766528.
XX 14-SEP-2000; 2000US-00661858.
XX (GEHO ) GEN HOSPITAL CORP.
XX Fishman JA;
XX WPI; 2004-689179/67.
XX N-PSDB; ADS73367.
XX New porcine retroviral polypeptide encoded by a nucleic acid, useful in
PT evaluating an immunosuppressive treatment for the ability to activate a
PT retrovirus, such as an endogenous porcine retrovirus.
XX Claim 8; Fig 1; 83pp; English.
XX The present invention relates to the swine retroviral polypeptides and
CC their encoding nucleic acids. The methods and compositions of the present
CC invention are useful for screening a cell or tissue, e.g. a heart, lung,
CC liver, bone marrow, kidney, brain cells, neural tissue, pancreas and
CC intestinal tissue xenograft, for the presence or expression of a swine or
CC miniature swine retrovirus or retroviral sequence. The invention is also
CC useful in evaluating an immunosuppressive treatment for the ability to
CC activate a retrovirus such as an endogenous porcine retrovirus. The
CC invention is also useful in gene therapy. The present sequence is the
CC swine retroviral (tsukuba-1) gag protein.
SQ Sequence 744 AA;
Alignment Scores:
Pred. No.: 1,03e-240 Length: 744
Score: 2813.00 Matches: 522
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 98.8% Indels: 0
DB: Gaps: 0
US-10-723-552-3_COPY_585_2156 (1-1572) x ADS73450 (1-744)
QY 1 ATGGACAGACGGTGACGACCCCTCTTAGTTTGACTCTCGACCATTTGACATGAAGTTAAA 60
Db 221 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValLys 240
QY 61 TCCAGGGCTCATAAATTTGTTCAGTTTAAAGAGGACCTTGGCAGACCTTCTGTCTC 120
Db 241 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysVal 260
QY 121 TCTGAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAAATTCGAGATT 180
Db 261 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 280
QY 181 ATCCTGGCTGTTAAAGCAGATTATTTTCAGACTGGACCCGGCTCTCATCCGATCAGAG 240
Db 281 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 300
QY 241 CCCTATATCCTTACGTGGCAAGATTGGCAGAGGATCTCCGCCATGGTGGTAAACCATGG 300
Db 301 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 320
QY 301 CTGAATAGCAAGAAAGCCAGGTCCCGGAATTTCTGGCTTTTGGAGAGAAAAACAACAC 360
```


Db 321 LeuAsnLysProArgLysProGlyProArgLysLeuAlaLeuGlyGluLysAsnLysHis 340
QY TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCACCGGCT 420
Db 341 SerAlaGluLysValLysProSerProHisLysLysProGluLysGluGluProProAla 360
QY TGGCCGGAAACCCCAATCTGTTCCCCACCCCTTATCTGGCAACAGGGTCCCGCAGGGGA 480
Db 361 TrpProGluProGlnSerValProProProProLysLeuAlaGlnGlyAlaAlaArgGly 380
QY CCCTTTGCGCCCTCTGAGCTCCGCGGTGGAGGACCTGTCGAGGACTCGAGCCGG 540
Db 381 PropheAlaProProGlyAlaProAlaValGluGlyProSerAlaGlyThrArgSerArg 400
QY AGGGCGCCACCCCGGAGCGACAGACGAGATCGGCATTTACCGCTGCGCAGCTACGGC 600
Db 401 ArgGlyAlaThrProGluArgThrAspGluLeuAlaThrLeuProLeuArgThrTyroGly 420
QY CCTCCACACCGGGGGGCAATTGACGCCCTCCAGTATTCGCCCTTTTCTTCGACAT 660
Db 421 ProProThrProGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 440
QY CTCTATAATTGGAAACTAAACCATCCCTTCTCGGAGGATCCCAACGGCTCACGGGG 720
Db 441 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 460
QY TTGCTGAGTCCCTTATGTTCTCTCACCGACTACTTGGGATGATTGTCAACAGCTGCTG 780
Db 461 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 480
QY 781 CAGACACTTTCAACACGAGGAGCGAGAGAAATTTCTATTAGAGGCTAGAAAATGTT 840
Db 481 GlnThrLeuPheThrThrGluGluArgGluArgLysLeuGluAlaArgLysAsnVal 500
QY 841 CCTGGGCGGAGCGGACCCACCGCTCAACAGCTCAAGGTAGGAGCTTGAAATCTATCGC 960
Db 501 ProGlyAlaAspGlyArgProThrArgLeuGlnAsnGluIleAspMetGlyPheProLeu 520
QY 901 ACTCGCCCGGTGGGACTCAACACGCTCAAGCTGAGGAGCTTGAAATCTATCGC 960
Db 521 ThrArgProGlyTrpAspPyrAsnThrAlaGluArgLysLeuLysIleTyrArg 540
QY 961 CAGGCTCTGGTGGGGTCTCCGGGGCCCTCAAGAGCGGCCACTAATTTGGCTTAAGTA 1020
Db 541 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 560
QY 1021 AGAAGTGATGAGGAGCGAATGAACCCCTCTGTTTCTTGAGAGCTCTTGAA 1080
Db 561 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuGlu 580
QY 1081 GCCTTCAGGGGTACACCCCTTTGATCCCACTCAGAGGCCCAAAAGCTCAGTGGCT 1140
Db 581 AlaPheArgTrpThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 600
QY 1141 TTGGCCCTTTATGAGACAGTCAAGCTCTGGATATTAGAAAGAGCTTCAGAGACTGGAAGG 1200
Db 601 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 620
QY 1201 TTACAGAGGCTGAGTTAGTGATCTAGTGAGGAGCGCAGAGAAAGTATTATTACAAAGG 1260
Db 621 LeuGlnGluAlaGluLeuArgAspLeuValLysGluAlaGluLysValTyrTyrLysArg 640
QY 1261 GAGACAGAGAAGAAAGGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 641 GluThrGluGluGluArgGluGluArgGluGluArgGluGluArgGluGluArgGlu 660
QY 1321 CGTATAAAGCGCAAGAGAGAAATTTGACTAAGATCTTGGCTGAGTGGTTGAGGAGAA 1380
Db 661 ArgAsnLysArgGlnGluLysAsnLeuThrLysIleLeuAlaValValGluGluLys 680
QY 1381 AGCAATACGGAAGAGAGAGATTTAGGAAATTTAGGTCAGGCCCTACACAGTCAAGG 1440

Db 681 SerAsnThrGluArgGluArgAspPheArgLysLysLeuArgSerGlyProArgGlnSerGly 700
QY 1441 AACCTGGGCAATAGCACCCCTCGCAAGACCAATGTGTATATTGTAAAGAAAGAGA 1500
Db 701 AsnLeuGlyAsnAlaThrProLeuAspLysAspGlnCysAlaTyrCysLysGluArgGly 720
QY 1501 CACTGGGCAAGGAATCTCCCAAGAGGGAAACAAAGACCAAGGATCTTAGCTCTAGAA 1560
Db 721 HistpAlaArgAsnCysProLysGlyAsnLysGlyProArgIleLeuAlaLeuGlu 740
QY 1561 GAGATAAAGAT 1572
Db 741 GluAspLysAsp 744
RESULT 6
ID AAO19009 standard; protein; 524 AA.
XX AAO19009;
AC AAO19009;
DT 07-NOV-2002 (first entry)
DE Porcine endogenous retrovirus type A gag.
KW PERV; protease; virucide; infection; xenotransplant; gag.
OS Porcine endogenous retrovirus.
PN DE10138528-Al.
XX 04-JUL-2002.
XX 06-AUG-2001; 2001DE-01038528.
XX 23-DEC-2000; 2000DE-01064897.
XX (UYMU-) UNIV MUECHEN MAXIMILIANS LUDWIG.
XX Von Der Helm K, Blusch JH, Seelmeir S;
DR WPI; 2002-609974/66.
DR N-PSDB; AAL49333.
XX New nucleic acid, useful for identifying specific inhibitors for
controlling infection following xenotransplantation, encodes porcine
endogenous retrovirus protease.
XX Disclosure; Page 14-16; 30pp; German.
XX The present invention relates to proteases and their coding sequences
from porcine endogenous retroviruses (PERVs) type A, B and C. The coding
sequences are useful in screening for inhibitors of proteases,
particularly where retroviral. The inhibitors are potentially useful for
treating or preventing PERV infections in humans, specifically those who
have received a pig organ transplant. The present sequence is a
retroviral gag protein
XX Sequence 524 AA;
Alignment Scores:
Pred. No.: 3,68e-231 Length: 524
Score: 2705.00 Matches: 497
Percent Similarity: 97.9% Conservative: 16
Best Local Similarity: 94.8% Mismatches: 11
Query Match: 95.0% Indels: 0
DB: Gaps: 5
US-10-723-552-3_COPY_585_2156 (1-1572) x AAO19009 (1-524)
QY 1 ATGGGACAGCGTGGAGCCCTCTAGTTTCTGACTCTCGACCTTGGACTGAGTTAAA 60
Db 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20

QY 61 TCCAGGCTCATATTTCTCAGTTTAAAGAGGACCTTGGCAGACTTTCTGTGTC 120
Db |||||
21 SerArgAlaHisAsnLeuSerValGlnValIysLysGlyProTrpGlnThrPheCysAla 40
QY 121 TCTGAATGCGGCACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAAATTTCTGAGATT 180
Db |||||
41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
QY 181 ATCTGGCTGTTAAAGAGATTATTTTTCAGACTGGACCGGCTCTCATCCCGATCAGGAG 240
Db |||||
61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
QY 241 CCTATATCTTACCTGTCAGAGATTGGCAGAGATCTCCGCCATGGTGTAAACCATGG 300
Db |||||
81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValIysProTrp 100
QY 301 CTGATAATAGCCAAAGAACCCAGGTCCCGAAATCTCTGGCTCTTGGAGAGAAAACAAACAC 360
Db |||||
101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGACGCCCGCT 420
Db |||||
121 SerAlaGluLysValGlyProSerProArgIleTyrProGluIleGluGluProProThr 140
QY 421 TGGCGGAACCCCAATCTGTTCCCGCCACCCCTTATCTGGCACAGGGTGCCGCGAGGGGA 480
Db |||||
141 TrpProGluProGlnProValProProProProTyrProAlaGlnGlyAlaValArgGly 160
QY 481 CCTTTGTCCTCTCTGGAGCTCCGCGGTGGAGGACCTGCTGCAGGGACTCGGAGCCGG 540
Db |||||
161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
QY 541 AGGGCGGCACCCCGAGGAGACAGACGAGATCGGCACATTACCGCTCGGCACGTACGGC 600
Db |||||
181 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200
QY 601 CCTCCACACCGGGGGCCAAATTCAGCCCTCCAGTATTCGCCCTTTCTTCGCAGAT 660
Db |||||
201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220
QY 661 CTCTATAATTGGAAACTAACCATCCCTTTCTCGAGAGATCCCAACGGCTCACGGGG 720
Db |||||
221 LeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
QY 721 TTGTTGGAGTCCCTTATGTTCTCTCACCAGCTACTTGGATGATGTCACAGCTGCTG 780
Db |||||
241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260
QY 781 CAGACACTCTTCACAAACGAGGAGGAGAGAGATTCTATTAGAGGCTAGAAAATGTT 840
Db |||||
261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVal 280
QY 841 CTGCGGCGCGAGCGGCGACCCACGCGGTTTGCAAAATGAGATTGACATGGGATTTCCCTTA 900
Db |||||
281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
QY 901 ACTGCGCCCGGTTGGGACTACAAACGCGCTGAAGCTAGGAGAGCTTGAAATCTATCGC 960
Db |||||
301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320
QY 961 CAGGCTCTGTTGGCGGGTCTCGGGCGCTCAAGACGCCCACTTAATTTGCTAAGGTA 1020
Db |||||
321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgProThrAsnLeuAlaLysVal 340
QY 1021 AGAGAAGTGATGACGGGACCGGAATGAACCCCTCTGTTTTTCTTGAGAGGCTCTTGAA 1080
Db |||||
341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
QY 1081 GCCTTCAGCGGTACACCCCTTTTGCATCCACCTCAGAGGCCCAAAAGCCTCAGTGCT 1140
Db |||||
361 AlaPheArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
QY 1141 TTGGCCTTTATAGGACAGTCAAGCTTGGATATTAGAAGAGACTTCAGAGACTGGAAGG 1200

Db ||||| LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
QY 1201 TTACAGAGGCTGAGTTACGTGATCTAGTGAAGGAGGAGAGAAAGTATATATACAAAAGG 1260
Db |||||
401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 420
QY 1261 GAGACAGNAGNAGAAAGGACAAAGAAAGAGAGAGAAAGAGAGAAAGGAGGAGAAAGA 1320
Db |||||
421 GluThrGluGluGluLysGluGlnArgLysGluLysGluArgGluGluArg 440
QY 1321 CGTAATAAACCAGCAAGAGAAGAAATTTGACTAAGATCTTGGCTGCAGTGGTTCAAGGGAAA 1380
Db |||||
441 ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLys 460
QY 1381 AGCAATACGGAAGAGAGAGAGATTATAGGAAATTAGGTAGGCGCTTAGACAGTCAGG 1440
Db |||||
461 SerSerArgGluArgGluArgAspPheArgIleArgSerGlyProArgGlnSerGly 480
QY 1441 AACCTGGCAATAGGACCCCACTCGACAGGACCAATGTGCATATTGTAAAGAAAGAGA 1500
Db |||||
481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500
QY 1501 CACTGGGCAAGAACTGCCCAAGAGGAAACAAAGAGCAACAGGATCCTAGCTCTAGAA 1560
Db |||||
501 HisTrpAlaArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
QY 1561 GAAGATAAAGAT 1572
Db |||||
521 GluAspLysAsp 524
RESULT 7
AAO19010
ID ARO19010 standard; protein; 524 AA.
XX
AC AAO19010;
XX
DT 07-NOV-2002 (first entry)
XX
DE Porcine endogenous retrovirus type B gag.
XX
KW PERV; protease; virucide; infection; xenotransplant; gag.
XX
OS Porcine endogenous retrovirus.
XX
FN DE10138528-Al.
XX
PD 04-JUL-2002.
XX
PF 06-AUG-2001; 2001DE-01038528.
XX
PR 23-DEC-2000; 2000DE-01064897.
XX
PA (UTMU-) UNIV MUENCHEN MAXIMILIANS LUDWIG.
XX
PI Von Der Helm K, Blusch JH, Seelmeir S;
XX
DR WPI; 2002-609974/66.
XX
PT N-PSDB; AAL49334.
XX
PT New nucleic acid, useful for identifying specific inhibitors for
controlling infection following xenotransplantation, encodes porcine
endogenous retrovirus protease.
XX
PS Disclosure; Page 16-17; 30pp; German.
XX
CC The present invention relates to proteases and their coding sequences
from porcine endogenous retroviruses (PERVs) type A, B and C. The coding
sequences are useful in screening for inhibitors of proteases, useful for
particulary where retroviral. The inhibitors are potentially useful for
treating or preventing PERV infections in humans, specifically those who
have received a pig organ transplant. The present sequence is a
retroviral gag protein

PF 13-DEC-1996; 96WO-US019680.
 XX
 XX
 XX 14-DEC-1995; 95US-00572645.
 XX
 XX (GBHO) GEN HOSPITAL CORP.
 XX
 XX Fishman JA;
 XX
 XX WPI; 1997-332804/30.
 DR N-PSDB; AAT74883.
 XX
 XX New nucleic acid from porcine retroviruses - used for detecting viruses
 PT in transplant or other tissue and for assessing risk of transmitting
 PT infection to graft recipient.
 XX
 XX Claim 16; Fig 2; 128pp; English.
 XX
 XX This sequence represents a putative viral GAG protein isolated from a
 CC defective porcine retrovirus found in PK-15 cells. This sequence and PCR
 CC fragments generated from it (see AAT74812-T74882) could be used to screen
 CC organs for porcine retroviruses prior to xenotransplantation.
 CC Transplantation can increase the likelihood of retroviral activation if
 CC intact and infectious proviruses are present. The porcine retroviral
 CC sequence can be used to generate probes to determine the level (e.g. copy
 CC number) of intact (i.e. potentially replicating) porcine provirus
 CC sequences in a strain of xenograft transplantation donors. It can be used
 CC to detect mutations, genetic lesions or viral recombinants and to
 CC determine the histological localisation of activated retrovirus. Using
 CC Polymerase Chain Reaction DNA Quantitation (PQ) on blood mononuclear
 CC cells, infectivity titration and susceptibility testing can be performed.
 CC Ultimately animal donors without intact porcine retroviral sequences or
 CC with a lower copy number of viral elements could be selected. (Updated on
 CC 27-AUG-2003 to correct OS field.)
 XX
 XX SQ Sequence 524 AA;

Alignment Scores:
 Pred. No.: 6,81e-231 Length: 524
 Score: 2702.00 Matches: 496
 Percent Similarity: 97.9% Conservative: 17
 Best Local Similarity: 94.7% Mismatches: 11
 Query Match: 94.9% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x AAW32091 (1-524)

Qy 1 ATGGGACAGCGTGACGACCCCTTTAGTTGACTCTCGACCAATGACGATGAAGTAA 60
 Db 1 MetGlyGlnThrValThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20

Qy 61 TCCAGGGCTCATATTTGTCAGTTTCAGTTTAAAGAGGACCTTGGCAGACTTCTGTCTC 120
 Db 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAla 40

Qy 121 TCTGAATGGCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTCGAGATT 180
 Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60

Qy 181 ATCTGGCTGTAAAGCAGATTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
 Db 61 IleLeuAlaValLysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80

Qy 241 CCTATATCTTACGTGGCAGATTGTCAGAGGATCTCCGCCATGGGTTAAACCATGG 300
 Db 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrp 100

Qy 301 CTGAATAGCCAAAGAACCCAGGTCCTCGAATTCGGCTCTTGGAGAGAAACAAACAC 360
 Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120

Qy 361 TCGGCTGAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGACCCCGGCT 420
 Db 121 SerAlaGluLysValGluProSerProArgIleTyrProGluIleGluGluProThr 140

Qy 421 TGGCCGGAAACCCCAATCTGTTCCCCACCCCTTATCTGGCACAGGGTGCCGCGAGGGA 480
 Db 141 TrpProGluProGlnProValProProProProTrpProAlaGlnGlyAlaValArgGly 160

Qy 481 CCCTTTGCCCTCTCTGGAGCTCCGGCGTGAGGAGCCTGCTGAGGAGACTCGGAGCGG 540
 Db 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180

Qy 541 AGGGCGCCACCCCGGAGCGGACAGAGATCCGACATTCACGCTCGCGCACGTACGGC 600
 Db 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGly 200

Qy 601 CTTCCACACACGGGGGCCAATTCAGCCCTCCAGTATTGGCCCTTTCTTCTGTCAGAT 660
 Db 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAsp 220

Qy 661 CTCTATAATTTGGAACCTAACCATCCCTTTCTCGGAGGATCCCAACGCTCACGGGG 720
 Db 221 LeuTyrAsnTrpLysThrAsnHisProPheSerGluAspProGlnArgLeuThrGly 240

Qy 721 TTGTGGAGTCCCTTATGTTCTCTCACAGCTACTTGGGATGATTTGTCACAGCTGCTG 780
 Db 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeu 260

Qy 781 CAGACACTCTTCACAAACCGAGGAGCGAGAGAAATTCATTAGAGGCTAGAAAAAATGTT 840
 Db 261 GlnThrLeuPheThrThrGluGluArgGluGluLeuLeuGluAlaLysLysAsnVal 280

Qy 841 CCTGGGGCCGACGGGCGACCCACGCGTTCGAAATGAGATTTGATGGGATTTCCCTTA 900
 Db 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300

Qy 901 ACTGCCCGGTTGGGACTACACACGCTGAAGTAGGAGGAGCTTGAAATCTATCGC 960
 Db 301 ThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArg 320

Qy 961 CAGGCTCTGTGGCGGCTCTCCGGGCGCCTCAAGACGGCCCTAATTTGGCTAAGGTA 1020
 Db 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340

Qy 1021 AGAGAAGTGTAGCAGGACCGAATGAACCCCTCTCTGTTTTTCTTGAGAGCTCTTGAA 1080
 Db 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360

Qy 1081 GCCTTCAGGGGTACACCCCTTTTGTATCCACCTCAGAGGCCCAAAAGCCCTCAGTGGCT 1140
 Db 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380

Qy 1141 TTGGCCTTTATAGCAGCTCAGCCCTTGATATTAGAAGAGCTTCAGAGACTCGAAGGG 1200
 Db 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400

Qy 1201 TTACAGAGGCTGAGTTACGTGATCTAGTGAAGAGGAGCAGAGAAAGTATATTACAAAAG 1260
 Db 401 LeuGlnGluAlaGluLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArg 420

Qy 1261 GAGACAGAAGAAAGGGAACAAAGAAAGAGAGAGAAAGAGAGAAAGGAGGAGAAAGA 1320
 Db 421 GluThrGluGluLysGluGlnArgLysGluLysGluArgGluGluArgGluGluArg 440

Qy 1321 CGTAATAACGGCAAGAGAGAAATTCGACTAAGATCTGGCTGAGTGGTTGAAGGGAAA 1380
 Db 441 ArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaIleValIleGluGlyLys 460

Qy 1381 AGCAATACGGAAGAGAGAGAGATTTTAGGAAAAATTAGGTCCAGCCCTTAGACGTACGGG 1440
 Db 461 SerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGly 480

Qy 1441 AACCTGGGCAATAGGACCCGCTCAGACGACCAATGTGCATATTGTAAAGAAAGAGGA 1500
 Db 481 AsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGly 500

QY 1501 CACTGGCAAGCACTGCCCAAGAGCGGACCAAGACCAAGGATCCTAGCTCTAGAA 1560
|||||
Db 501 HisTrpAlaArgAsnCyseProLysGlyAsnLysGlyProLysValLeuAlaLeuGlu 520
|||||
QY 1561 GAAAGATAAGAT 1572
|||||
Db 521 GluAspLysAsp 524
|||||
RESULT 9
AAB73282
ID AAB73282 standard; protein; 524 AA.
XX
AC AAB73282;
XX
DT 23-MAY-2001 (first entry)
XX
DE Defective retroviral genome protein #1 isolated from PK-15 cell line.
XX
KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line.
XX
OS Unidentified.
XX
PN US6190861-B1.
XX
PD 20-FEB-2001.
XX
PF 13-DEC-1996; 96US-00766528.
XX
PR 14-DEC-1995; 95US-00572645.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI; 2001-256211/26.
DR N-PSDB; AAF77726.
XX
PT Assessing risk of endogenous retroviruses in clinical practice and in
PT xenotransplantation, comprises using probe sequences derived from swine
PT or miniature swine retroviral genome.
XX
PS Disclosure; Fig 2; 127pp; English.
XX
CC The present invention relates to a method for screening a cell or tissue
CC for the presence or expression of a retrovirus (RV), comprising
CC contacting a target nucleic acid from the cell or tissue with a second
CC nucleic acid from the present invention (e.g. AAF77726 or a fragment
CC thereof). The method is useful for RV detection and to assess graft
CC transplantation risk. Screening of animals allows the elimination of
CC donors with active replication of known viruses. Inactive proviruses can
CC be detected and inactivated, allowing identification and elimination of
CC potential human pathogens derived from swine in a manner not possible in
CC the outbred human organ donor population and is important to the
CC development of human xenotransplantation
XX
SQ Sequence 524 AA;
Alignment Scores:
Pred. No.: 6.81e-231 Length: 524
Score: 2702.00 Matches: 496
Percent Similarity: 97.9% Conservative: 17
Best Local Similarity: 94.7% Mismatches: 11
Query Match: 94.9% Indels: 0
DB: 4 Gaps: 0
US-10-723-552-3_COPY_585_2156 (1-1572) x AAB73282 (1-524)
QY 1 ATGGGACAGACGGTGACGACCCCTCTAGTTTGACTCTCGACCACTTGGACTGAAGTAAA' 60
|||||
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20
|||||
QY 61 TCCAGGCTCATAAATTTGTCAAGTTCAAGTTAAGAGGACCTTGGCAGACTTCTGTGTC 120
|||||

Db 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAla 40
QY 121 TCTGAATGCCGACATTCGATGTTGGATGGCCATCAGAGGGAGACCTTTAAATTCAGATT 180
|||||
Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluLeu 60
|||||
QY 181 ATCTGGCTGTTAAAGCAGTTATTTTCAGACTGGACCCCGCTCTCATCCGATCAGAG 240
|||||
Db 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
|||||
QY 241 CCTATATCTTACGTGTCAGATTTCGCACAGGATCCTCCGATCAGTTGTTAAACCATCG 300
|||||
Db 81 ProTrpIleLeuThrTrpGlnAspLeuAlaGluAspProProTrpValLysProTrp 100
|||||
QY 301 CTGAATAAGCCAAAGACCCAGGTCCCGCAATTCGTGCTCTGGAGAGAAAAACACAC 360
|||||
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHis 120
|||||
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATATCTACCCCGAGATGAGGACCCAGCGCT 420
|||||
Db 121 SerAlaGluLysValGluProSerProArgIleTrpProGluIleGluGluProProThr 140
|||||
QY 421 TGGCCGGAACCCCAATCTGTTCCGCCACCCCTTATCTGGCACAGGTCGCCGAGGGA 480
|||||
Db 141 TrpProGluProGlnProValProProProTrpProAlaGlnGlyAlaValArgGly 160
|||||
QY 481 CCTTTTGGCTCTCTGGAGCTCCCGCGGTGGAGGACCTGCTGAGGAGCTCGAGCGCG 540
|||||
Db 161 ProSerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArg 180
|||||
QY 541 AGGGCGCCACCCCGAGCGGACAGACAGATCCCGCATTTACCGCTGCGCACCTACCGC 600
|||||
Db 181 ArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrgly 200
|||||
QY 601 CCTCCACACCGGGGGGCAATTCAGCCCTCCAGTATTCGCCCTTTCTCTTCGCAGAT 660
|||||
Db 201 ProProMetProGlyGlyGlnLeuGlnProLeuGlnTyTrpProPheSerSerAlaAsp 220
|||||
QY 661 CTCTATATTTGGAATCTAAACATCCCTTTCTCGAGGATCCCAACGCTCACGCGG 720
|||||
Db 221 LeuTyraAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGly 240
|||||
QY 721 TTGCTGGAGTCCCTTATGTTCTCTCACAGCTACTTGGGATGATTTGTCAACAGCTGCTG 780
|||||
Db 241 LeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnLeuLeu 260
|||||
QY 781 CAGACATCTTCAACACCGAGGCGGAGAGAGAAATCTATTAGAGCTAGAAAAATGTT 840
|||||
Db 261 GlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaLysLysVal 280
|||||
QY 841 CCTGGGCGCAGCGGCGACCCAGCGGTTCGAAATCAGATTGATGAGGATTTCCCTTA 900
|||||
Db 281 ProGlyAlaAspGlyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeu 300
|||||
QY 901 ACTCGCCCGGTTGGGACTACACACCGCTGAAGGTAGGAGAGCTTGAATAATCTATCGC 960
|||||
Db 301 ThrArgProGlyTrpAspTrpAsnThrAlaGluGlyArgGluSerLeuLysIleTyArg 320
|||||
QY 961 CAGGCTCTGTGGCGGTCTCCGGGCGCTCAAGACGGCCCACTAATTTGGCTAAGATA 1020
|||||
Db 321 GlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVal 340
|||||
QY 1021 AGAGAGTGTATGAGGAGCGCAATGACCCCTCTGTTTCTTTGAGAGGCTCTTGGAA 1080
|||||
Db 341 ArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGlu 360
|||||
QY 1081 GCCTTCAGGCGGTACACCCCTTTTGATCCCACTCCAGAGGCCCAAAAGCCCTCAGTGCT 1140
|||||
Db 361 AlaPheArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAla 380
|||||
QY 1141 TTGGCTTTTATAGACAGTCAAGCTTGGATATTAGAAAGAGCTTCAGAGACTGGAAGG 1200
|||||
Db 381 LeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGly 400
|||||


```
Db 180 gArgGlyAlaThrProGluAArgThrAspGluLeuAlaIleLeuProLeuAArgThrTyrGI 200
Qy 600 CCCTCCACACCGGGGGCCCAATTGCACCCCTCCAGTATTGGCCCTTTCTTCTGCAGA 659
Db 200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAs 220
Qy 660 TCTCTATATTTGGAACTACCATCCCTTCTCCGAGGATCCCAACGCTCACGGG 719
Db 220 pLeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnAArgLeuThrGI 240
Qy 720 GTTGCTGAGTCCCTTATGTTCTCTCACGAGCTACTTGGGATGATTGTCAACAGCTGCT 779
Db 240 yLeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLe 260
Qy 780 GCAGACACTCTTCAACCCGAGGAGCGAGAGAAATCTATTAGAGGCTAGAAAATGT 839
Db 260 uGlnThrLeuPheThrThrGluAArgGluAArgIleLeuLeuGluAAlaAArgLysAsnVa 280
Qy 840 TCCTGGGCGCCAGCGGCGACCCACCGCTTGCAAAATGAGATTGACATGGGATTCCTT 899
Db 280 lProGlyAlaAspGlyAArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe 300
Qy 900 AACTCGCCCGCTTGGGACTCAACACCGCTGAAAGTAGGAGACTTGAATAATCTATCG 959
Db 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyAArgGluSerLeuLysIleTyrAr 320
Qy 960 CCAGGCTCTGTGGCGGTCTCCGGGCGCTCAAGACGGCCCACTAATTTGGCTAAGGT 1019
Db 320 gGlnAlaLeuValAlaGlyLeuAArgGlyAlaSerArgArgProThrAsnLeuAlaLysVa 340
Qy 1020 AAGAGAGTGTAGCAGGACCGAATGAACCCCTCTGTTTCTTGAGAGCTCTTGA 1079
Db 340 lArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluAArgLeuMetGI 360
Qy 1080 AGCCTTCAGGCGGTACACCCCTTTTGATCCCACTCAGAGGCGCCAAAAGCTCAGTGGC 1139
Db 360 uAlaPheAArgArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAl 380
Qy 1140 TTTGGCTTTTATAGACAGTCAAGCTCTGGATATTAGAAAGAGCTTCAGAGACTGGAAG 1199
Db 380 aLeuAlaPheIleGlyGlnSerAlaLeuAspIleAArgLysLysLeuGlnAArgLeuGluGI 400
Qy 1200 GTTACAGAGGCTGATGTACGTGATCTAGTGAAGGACGACGAGAAAGTATATTACAAAG 1259
Db 400 yLeuGlnGluAlaGluLeuAArgAspLeuValAArgGluAlaGluLysValTyrTyrArgAr 420
Qy 1260 GGAGACAGAAGAAGAAAGGAAACAAAGAAAGAGAGAGAAAGAGAGAGAAAGGAGAAAG 1319
Db 420 gGluThrGluGluGluLysGluGlnAArgLysGluLysGluAArgGluAArgGluGluAr 440
Qy 1320 ACGTAATAAACCGCAAGAGAAGAAATTTGACTAAGATCTTGGCTGCAGTGGTTGAAGGAA 1379
Db 440 gArgAspArgArgGlnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLy 460
Qy 1380 AAGCAATACGGAAGAGAGAGAGATTTAGGAAATTAGTTCAGGCCCTAGACAGTCAAG 1439
Db 460 sSerSerArgGluAArgGluAArgAspPheAArgLysIleAArgSerGlyProAArgGlnSerGI 480
Qy 1440 GAACCTGGGCAATAGGACCCACTCGACAAAGGACCAATGTGCATATTGTAAAGAAAGAGG 1499
Db 480 yAsnLeuGlyAsnAArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGI 500
Qy 1500 ACATGGGCAAGGAATCGCCCCAAGAGGGAACAAAGGACCAAGGATCTAGCTCTAGA 1559
Db 500 yHisTrpAlaAArgAsnCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGI 520
Qy 1560 AGAAGATAAAGAT 1572
Db 520 uGluAspLysAsp 524
RESULT 12
AAB70629
```

```
ID AAB70629 standard; protein; 524 AA.
XX
AC AAB70629;
XX
DT 15-MAY-2001 (first entry)
XX
DE Porcine endogenous retrovirus gag amino acid sequence.
XX
KW Porcine endogenous retrovirus; PoERV; gag; env; detection; antigenic;
KW immunogenic; antiviral; vaccine; antiserum; viral infection.
XX
OS Porcine endogenous retrovirus.
XX
PN W0200112816-A2.
XX
PD 22-FEB-2001.
XX
PF 16-AUG-2000; 2000WO-GB003159.
XX
PR 18-AUG-1999; 99GB-00019604.
XX
PA (QUIP-) QUIP TECHNOLOGY LTD.
XX
PI Galbraith D, Kelly H, Smith K;
XX
DR WPI, 2001-211223/21.
XX
PT New porcine endogenous retrovirus polypeptide fragment with retroviral
PT specific antigenic or immunogenic activity, for detection of retroviral
PT antibodies in a sample and in therapy or diagnosis.
XX
PS Claim 3; Fig 1; 43pp; English.
XX
CC The present invention describes a porcine endogenous retrovirus (PoERV)
CC polypeptide fragment (I) which has PoERV specific antigenic or
CC immunogenic activity. Also described are: (1) an antiserum (II) specific
CC to (I); (2) a PoERV specific antibody (III) or its fragment raised
CC against (I); (3) use of (I) in detecting (III) in a sample; (4) use of
CC (III) in detecting PoERV in a sample; (5) use of (I) or (III) in therapy
CC or diagnosis; and (6) an assay kit (IV) comprising (I) or (III) for use
CC in detection of PoERV in a sample. (I) has antiviral activity, and can be
CC used in vaccine production. (I) is useful in the detection of PoERV
CC antibodies (III) in a sample and (III) is useful in the detection of
CC PoERV in a sample. (I) and (III) are useful in therapy or diagnosis.
CC Antiserum (II) to (I) and kits (IV) comprising (I) or (III) are useful in
CC detection of specific PoERV types. Viral infection can be monitored and
CC by using (I) or (III). The present sequence represents the specifically
CC claimed PoERV gag amino acid sequence from the present invention
XX
SQ Sequence 524 AA;
```

```
Alignment Scores:
Pred. No.: 2,21e-225 Length: 524
Score: 2640.00 Matches: 495
Percent Similarity: 97.3% Conservative: 16
Best Local Similarity: 94.3% Mismatches: 13
Query Match: 92.7% Indels: 2
DB: 4 Gaps: 0

US-10-723-552-3_COPY_585_2156 (1-1572) x AAB70629 (1-524)
Qy 1 ATGGGACAGACGGTGACGACCCCTCTTAGTTTGACTCTCGACCATTTGGACTGAAGTTAAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpThrGluValArg 20
Qy 61 TCCAGGGCTCATAAATTTGTTCAGTTTAAAGAGGGACCTTGGCAGACTTTCTGTGTC 120
Db 21 SerArgAlaHisAsnLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAla 40
Qy 121 TCTGAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTTCTGATTT 180
Db 41 SerGluTrpProThrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIle 60
```

QY 181 ATCTGGCTGTTAAAGCAGTTATTTTCAGACTGACCCGGCTCTCATCCGATCAGAG 240
Db 61 IleLeuAlaValLysAlaIlePheGlnThrGlyProGlySerHisProAspGlnGlu 80
QY 241 CCCTATATCCCTACCTGCAAGATTGGCAGAGGATCCCGCATGGGTAAACCATGG 300
Db 81 ProTyrIleLeuThrTrpGlnAspLeuAlaGlnAspProProTrpValIlyProTrp 100
QY 301 CTGAATAAGCCAAAGAACCCAGGTCCTCCGAAATTCCTGGCTCTTGGAGAGAAAAACAAACAC 360
Db 101 LeuAsnLysProArgLysProGlyProArgIleLeuAlaLeuGlyLysAsnLysHis 120
QY 361 TCGGCTGAAAAAGTCAACCCCTC-TCCTCATATCTACCCCGAGATTGAGAGCACCGCC 419
Db 121 SerAlaGluLysValGluProSerSerTyrLeuProArgAspArgGlyAlaAlaAsp 140
QY 420 TTGGCCGGAACCCCAATCTGTTCCCAACCCCTTATCTGACAGAGGTGCGGAGGGG 479
Db 141 LeuAlaGlyThrProThrCysSerProThrProLeuSerSerThrGlyCysGluGly 160
QY 480 ACCCTTTGCCCTCTCGAGCTCCGGCGTGGAGGACCTCTGTCAGGAGCTCGAGCGC 539
Db 161 Thr-SerAlaProProGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerAr 180
QY 540 GAGGGGCCCAACCCCGAGCGGACAGCAGATCGGACATTACCGCTGCCACGTACGG 599
Db 180 GArgGlyAlaThrProGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGl 200
QY 600 CCCTCCCAACACCGGGGCCAATTCGACGCCCTCCAGTATTGGCCCTTTCTTCGAGA 659
Db 200 yProProMetProGlyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAs 220
QY 660 TCTCTATAATTGAAAACTAACCATCCCTCTTCGAGGATCCCAACCCCTCACGGG 719
Db 220 pLeuTyrAsnTrpLysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGl 240
QY 720 GTTGGTGAGTCCCTTATGTTCTCTCACAGCTTACTTGGGATGATGTCAACAGTCT 779
Db 240 yLeuValGluSerLeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeu 260
QY 780 GCACACTCTTCAACCGGAGGAGGAGAGATTTCTATTAGAGCTAGAAAAATGT 839
Db 260 uGlnThrLeuPheThrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnVa 280
QY 840 TCCTGGGGCCGACCGGCGCTTCAAAATGAGATTGACATGGATTTCCCTT 899
Db 280 lProGlyAlaAspLysArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLe 300
QY 900 AACTCGCCCGGTTGGGACTACACACGGCTGAAGGTAGGAGAGCTTGAATAATCTATCG 959
Db 300 uThrArgProGlyTrpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrAr 320
QY 960 CCAGGCTCTGTGGGGCTCTCCGGGGCTCAAGACGCCCTCAATTTGGCTAGGT 1019
Db 320 gGlnAlaLeuValAlaGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysVa 340
QY 1020 AAGAGAAAGTGTATGACGAGCCGAATGAACCCCTCTGTGTTTCTTCAGAGGCTCTTGA 1079
Db 340 lArgGluValMetGlnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGl 360
QY 1080 AGCTTCAGGGGTACACCCCTTTTGATCCCACTCAGAGCCCAAAAGCCTCAGTGGC 1139
Db 360 uAlaPheArgPheThrProPheAspProThrSerGluAlaGlnLysAlaSerValAl 380
QY 1140 TTTGGCTTTATAGCAGTCAGCTGGATATTAGAAAGAGCTTCAGAGCTGGAGG 1199
Db 380 aLeuAlaPheIleGlyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGlu 400
QY 1200 GTTACAGAGGCTGAGTTACTGTATCTAGTGAAGGCGCAGAGAAAGTATTACAAAG 1259
Db 400 yLeuGlnGluAlaGluLeuArgPheLeuValArgGluAlaGluLysValTyrTyrArgAr 420
QY 1260 GGAGACAGAAAGAAAGGGAACAAAGAAAGAGAGAGAAAGAGAGAAAGGAGGAAAG 1319

Db 420 gGluThrGluGluGlyLysGluGlnArgLysGluLysGluArgGluGluArgGluGluAr 440
QY 1320 ACGTAATAAAGCGCAAGAGAGAAATTTCACTAAGATCTTGCTCGAGTGGTTGAAGGAA 1379
Db 440 gArgAspArgGlnGlnLysAsnLeuThrLysIleLeuAlaAlaValAlaGluGlyLy 460
QY 1380 AAGCAATACGAAAGAGAGAGATTTTAGGAAAATTAGGTGAGGAGGATCCAGCTCAGG 1439
Db 460 sSerSerArgGluArgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGl 480
QY 1440 GAACCTGGGCAATAGGACCCCTCGACAGGACCAATGTGCATATTGTAAAGAAAGG 1499
Db 480 yAsnLeuGlyAsnArgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGl 500
QY 1500 ACACCTGGCAGGAACTGCCCAAGAGGGAACCAAGGACCAAGGATCCCTAGCTCTAGA 1559
Db 500 yHisTrpAlaArgAsnCysProLysGlyAsnLysGlyProLysValLeuAlaLeuGl 520
QY 1560 AGAAGATAAAGAT 1572
Db 520 uGluAspLysAsp 524
RESULT 13
AAW81571
ID AAW81571 standard; protein; 522 AA.
XX AC AAW81571;
XX DT 01-MAR-1999 (first entry)
XX DE Mus dunni endogenous virus Gag protein.
XX KW MDEV; retrovirus; packaging cell line; gene transfer; gene therapy;
XX KW vector; Gag protein.
XX OS Mus dunni endogenous virus.
XX PN WO9850538-A1.
XX PD 12-NOV-1998.
XX PF 08-MAY-1998; 98WO-US009452.
XX PR 09-MAY-1997; 97US-0046140P.
XX PR 08-MAY-1998; 98US-00075272.
XX PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX PI Miller AD, Wolgamot G, Bonham L;
XX DR WPI; 1999-034718/03.
XX DR N-PSDB; AAV69750.
XX PT New retroviral packaging cells - containing Mus dunni endogenous virus
XX PT sequences to target cells, retrovirus gag and pol genes and a
XX PT heterologous gene of interest.
XX PS Disclosure; Page 67-68; 85pp; English.
XX
CC This is the amino acid sequence of the Gag protein of Mus dunni
CC endogenous virus (MDEV), as deduced from the MDEV nucleotide sequence
CC (see AAV69750). A cultured packaging cell is claimed which produces a
CC replication-defective retroviral vector (RDRV) particle, where the
CC packaging cell is a vertebrate cell capable of expressing and assembling
CC retroviral proteins, comprising: (a) a first vector encoding a retroviral
CC envelope protein having amino acid residues MDEV that direct binding of
CC the retroviral particle to MDEV retroviral receptors on a target cell;
CC and (b) a second vector encoding retrovirus Gag and Pol proteins, where
CC upon expression of the vectors in the packaging cell in the presence of a
CC vector having a sequence of a heterologous gene of interest, a
CC replication-defective retroviral particle is produced that binds to MDEV
CC receptors of target cells. Also claimed are: (1) a cultured packaging

cell for producing a RDRV particle; (2) methods for producing a RDRV particle comprising a heterologous gene of interest; (3) cultured packaging cell line PD223; and (4) a RDRV produced by a method as in (2). The MDEV receptor is present on a variety of cells rendering MDEV pseudovirion packaging cells useful in methods of mammalian and particularly human gene transfer for gene therapy. The MDEV packaging cells are a stable and reproducible source of retroviral particles. Clones may be isolated from these populations that produce high titre virus. The packaging cell lines may be selected and cloned for other desirable properties, such as stability of in vivo growth, lack of production of helper virus, lack of reinfection by viral particles packaged in the cell, stability from genetic rearrangement and recombinational events, resistance to complement lysis, and improved ability to infect cells from higher mammals

XX
SQ Sequence 522 AA;

Alignment Scores:
Pred. No.: 9, 08e-145 Length: 522
Score: 1733.00 Matches: 340
Percent Similarity: 73.3% Conservative: 61
Best Local Similarity: 62.2% Mismatches: 96
Query Match: 60.8% Indels: 50
DB: 2 Gaps: 12

US-10-723-552-3_copy_585_2156 (1-1572) x AAW81571 (1-522)

QY 1 ATGGGACAGCGGTGACGACCCCTCTTAGTTGACTCTCGACCAATTGACGTGAAGTTAAA 60
DB 1 MetGlyGlnSerValSerThrProLeuSerLeuThrLeuGluHisTrpLysGluValLys 20
QY 61 TCCAGGCTCATATTTCTAGTTTCAAGTTAAGAGGACCTTGGCAGACTTCTGTCTC 120
DB 21 IleArgAlaHisAsnGlnSerValGluValArgLysGlyProTrpGlnThrPheCysAla 40
QY 121 TCTGAATGCGCGACATTCGATGTTGGATGGCCATCAGAGGGGACCTTAAATTTCTGAGATT 180
DB 41 SerGluTrpProThrPheGlyValGlyTrpProGluGlyAlaPheAspLeuSerLeu 60
QY 181 ATCTGGCTGTTAAGCAGTATTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
DB 61 IleAlaAlaValArgIleValPheGlnGlu---GluGlyGlyHisProAspGlnIle 79
QY 241 CCTATATCTTACGTGACAGATTGGCAGAGGATCTCCGCCATGGTGGTTAAACCATGG 300
DB 80 ProTyrIleValThrTrpGlnSerLeuValGlnPheProProSerTrpValIleProTrp 99
QY 301 CTGAATAAGCCAAGAACCCAGGTCCTCCGAAATTCGGCTCTTGGGAGAGAAAAACAACAC 360
DB 100 -----ThrProAsnProSerLysLeuThrValAlaValAlaGlnSerAsp 114
QY 361 TCGGCTGAAAAGTCAAGCCCTCT-----CCTCATATCTACCCCGAGATTGAG----- 408
DB 115 AlaAlaGluLysSerGlyProSerAlaProProLysIleTyrProGluIleAspAspLeu 134
QY 409 -----GAGCCACCGCTTGGCGGAGACCCCAATCTGTTCCC-----CCA 447
DB 135 LeuTrpMetAspSerGlnProProTrpProLeuProGlnGlnProProAlaAlaAla 154
QY 448 CCCCCTTATCTGGCACAGGGTCCCGGAGGGGACCCCTTGGCCCTCTCTGGAGCTCCGCG 507
DB 155 ProProValAlaAla-----ProGlnProGluProThrAlaSerGly 168
QY 508 GTGAGGACCTGTGTCAGAGGACTCGAGCCGGAGGGCGCCACCCCGAGCGG----- 561
DB 169 AlaGlnGlyProAlaGlyGlyThrArgSerArgGlyArgSerProAlaGluGluGly 188
QY 562 ---ACAGCAGAGATCGCACATTACCGCTGGCAGCTAC---GGCCCTCCACACCGGG 615
DB 189 GlyProAspSerThrValAlaLeuProLeuArgAlaHisValGlyGlyProThrProGly 208
QY 616 ---GGCCAATGACGCCCTCCAGTATTGGCCCTTTCTCTGTCAGATCTCTATAATTGG 672

DB 209 ProAsnAspLeuIleProLeuGlnTyrTrpProPheSerSerSerAspLeuTyrAsnTrp 228
QY 673 AAAACTAACCATCCCTTCTCGAGAGATCCCAACGCCCTCAGGGGTGGTGGAGTCC 732
DB 229 LysThrAsnHisProProPheSerGluAsnProSerGlyLeuThrGlyLeuGluSer 248
QY 733 CTTATGTTCTCTCACAGCCTTACTTGGATGATTGTCAACAGCTGCTGCACACACTCTTC 792
DB 249 LeuMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnValLeuPhe 268
QY 793 ACAACGAGGAGCGAGAGAGATTTCTATTAGAGCTAGAAAAATGTTCTCGGGCGCAG 852
DB 269 ThrThrGluGluArgGluArgIleLeuMetGluAlaArgLysAsnValLeuGlyGluAsp 288
QY 853 GGGGACCCACCGCGTTCGAAAAATGAGATTGACATGGGATTTCCCTTAACCTCGCCCGGT 912
DB 289 GlyThrProThrAlaLeuProAsnLeuValAspGluAlaPheProLeuAsnArgProAsn 308
QY 913 TGGGACTACACACGCGCTGAAGGTAGGAGAGCTTGAATAATCTATCGCAGGCTCTGGTG 972
DB 309 TrpAspTyrAsnThrAlaGluGlyArgGlyArgLeuLeuValTyrArgThrLeuVal 328
QY 973 GCGGGTCTCCGGGCGCGCTCAAGACGCGCCACTAATTTGGCTAAGTAAGAGAAGTATG 1032
DB 329 AlaGlyLeuArgGlyAlaAlaArgArgProThrAsnLeuAlaLysValArgGluValLeu 348
QY 1033 CAGGACCGAATGAACCCCTCTGTTTCTTGAGAGGCTCTTGAAGCTTTCAGGCGG 1092
DB 349 GlnGlyGlnThrGluProProSerValPheLeuGluArgLeuMetGluAlaTyrArgArg 368
QY 1093 TACACCCCTTTGATCCACCTCAGAGCGCCCAAAAGCCTCAGTGGCTTTGGCCTTTATA 1152
DB 369 TyrThrProPheAspProSerSerGlyGlnLysAlaAlaValAlaMetAlaPheIle 388
QY 1153 GGACAGTCAGCTTGGATATTAGAAAGAAGCTTCAGAGACTGGAAGGTTACAGAGGCT 1212
DB 389 GlyGlnSerAlaProAspIleLysLysLysLeuGlnArgLeuGluGlyLeuGlnAspTyr 408
QY 1213 GAGTTAGTGTAGTGAAGGAGCAGAGAAAGTATTATTAACAAGGGAGACAGAGAA 1272
DB 409 ThrLeuGlnAspLeuValLysGluAlaGluLysValTyrHisLysArgGluThrGluGlu 428
QY 1273 GAAAGGAAACAAAGAAAG 1332
DB 429 GluArgGlnGluArgGlyLysGluValGluGluArgGluAsnArgArgAspArgArg 448
QY 1333 CAAGAGAAGATTTGACTAAGATCTTGCTCGAGTGTGTTGAAGGAAAGCAATACGAA 1392
DB 449 GlnGluArgAsnLeuSerLysIleLeuAlaAlaValIleAsnAspArg----- 464
QY 1393 AGAGAGAGAGATTTTAGAAAAATTAGTCAAGCCCTAGACAGTCAAGGAACTCGGCAAT 1452
DB 465 -----GlnSerGluLysGlyArgThrGlyPheLeuGlyAsn 476
QY 1453 AGG-----ACCCCACTCGAACAGGACCAATGTGCATAT 1485
DB 477 ArgAlaValLysProProGlyGlyArgLysThrProLeuGluLysAspGlnCysAlaPhe 496
QY 1486 TGTAAAGAAAGAGACACTGGGCAAGGAACTGCCCAAGAGGAGGAAACAAAGGACCAAG 1545
DB 497 CysLysGluLysGlyHisTrpAlaLysAspCysProLysLys---ArgArgGlnPheLys 515
QY 1546 ATCTAGCTCTAGAAGAAGAT 1566
DB 516 ValLeuThrLeuGluAspAsp 522
RESULT 14
ID AAW81570 standard; protein; 622 AA.
XX AAW81570;
XX
DT 01-MAR-1999 (first entry)

XX	Mus dunni endogenous virus Gag protein (glycosylated).
DE	MDEV; retrovirus; packaging cell line; gene transfer; gene therapy;
XX	vector; Gag protein.
KW	
KX	
XX	
OS	Mus dunni endogenous virus.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 1
FT	/note= "encoded by CTG"
XX	
PN	WO9850538-A1.
XX	
PD	12-NOV-1998.
XX	
PF	08-MAY-1998; 98WO-US009452.
XX	
PR	09-MAY-1997; 97US-0046140P.
PR	08-MAY-1998; 98US-00075272.
XX	
PA	(HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX	
PI	Miller AD, Wolgamot G, Bonham L;
XX	
DR	WPI; 1999-034718/03.
DR	N-PSDB; AAV69750.
XX	
PT	New retroviral packaging cells - containing Mus dunni endogenous virus
PT	sequences to target cells, retrovirus gag and pol genes and a
PT	heterologous gene of interest.
XX	
PS	Disclosure; Page 65-66; 85pp; English.
XX	
CC	This is the amino acid sequence of the glycosylated Gag protein of Mus
CC	dunni endogenous virus (MDEV), as deduced from the MDEV nucleotide
CC	sequence (see AAV69750). A cultured packaging cell is claimed which
CC	produces a replication-defective retroviral vector (RDRV) particle, where
CC	the packaging cell is a vertebrate cell capable of expressing and
CC	assembling retroviral proteins, comprising: (a) a first vector encoding a
CC	retroviral envelope protein having amino acid residues MDEV that direct
CC	binding of the retroviral particle to MDEV retroviral receptors on a
CC	target cell; and (b) a second vector encoding retrovirus Gag and Pol
CC	proteins, where upon expression of the vectors in the packaging cell in
CC	the presence of a vector having a sequence of a heterologous gene of
CC	interest, a replication-defective retroviral particle is produced that
CC	binds to MDEV receptors of target cells. Also claimed are: (1) a cultured
CC	packaging cell for producing a RDRV particle; (2) methods for producing a
CC	RDRV particle comprising a heterologous gene of interest; (3) cultured
CC	packaging cell line PD223; and (4) a RDRV produced by a method as in (2).
CC	The MDEV receptor is present on a variety of cells rendering MDEV
CC	pseudotype packaging cells useful in methods of mammalian and
CC	particularly human gene transfer for gene therapy. The MDEV packaging
CC	cells are a stable and reproducible source of retroviral particles.
CC	Clones may be isolated from these populations that produce high titre
CC	virus. The packaging cell lines may be selected and cloned for other
CC	desirable properties, such as stability of in vivo growth, lack of
CC	production of helper virus, lack of reinfection by viral particles
CC	packaged in the cell, stability from genetic rearrangement and
CC	recombinational events, resistance to complement lysis, and improved
CC	ability to infect cells from higher mammals
XX	
SQ	Sequence 622 AA;
	Alignment Scores:
Pred. No.:	9.59e-145 Length: 622
Score:	1733.00 Matches: 340
Percent Similarity:	73.3% Conservative: 61
Best Local Similarity:	62.2% Mismatches: 96
Query Match:	60.8% Indels: 50
DB:	2 Gaps: 12
	US-10-723-552-3 COPY 585 2156 (1-1572) x AAW81570 (1-622)

```
QY 1033 CAGGACCGAATGAACCCCTCTGCTTTCTTGAGAGGCTCTTGAGAGCTTCAGGCGG 1092
Db 449 GlnGlyGlnThrGluProProSerValPheLeuGluArgLeuMetGluAlaItyrArgArg 468
QY 1093 TACACCCCTTTTGATCCACCTTCAGAGCCCAAAAGCCTCAGTGGCTTTGGCTTTTATA 1152
Db 469 TyrThrProPheAspProSerSerGluGlyGlnLysAlaAlaValAlaMetAlaPheIle 488
QY 1153 GCAGAGTCAGCTTGATATTAGAAAGCTTCAGAGACTGGAGGTTTACAGAGGCT 1212
Db 489 GlyGlnSerAlaProAspIleLysLysLysLeuGlnArgLeuGluGlyLeuGlnAspTyr 508
QY 1213 GAGTTACTGCTAGTCAAGAGCGCAGAGAAAGTATATTACAAAGGGAGACAGAGAA 1272
Db 509 ThrLeuGlnAspLeuValLysGluAlaGlnLysValTyrHisLysArgGluThrGluGlu 528
QY 1273 GAAAGGGAACAAAGAAAGAGAGAAAGAGAGAAAGGAGAAAGCGTAATAAAGCG 1332
Db 529 GluArgGlnGluArgGluLysLysGluValGluGluArgGluAsnArgArgAspArg 548
QY 1333 CAAGAGAAAGATTGACTTAAGATCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
Db 549 GlnGluArgAsnLeuSerLysIleLeuAlaValIleAsnAspArg----- 564
QY 1393 AGAGAGAGAGATTTAGGAAATAGGTCAGGCCCTAGACAGCTCAGGGAACCTGGCAAT 1452
Db 565 -----GlnSerGluLysGlyArgThrGlyPheLeuGlyAsn 576
QY 1453 AGG-----ACCCCACTCGACAGGCAATGTGCATAT 1485
Db 577 ArgAlaValLysProProGlyGlyArgLysThrProLeuGluLysAspGlnCysAlaPhe 596
QY 1486 TGTAAGAAAGAGCACTGGGCAAGAACTGCCCCCAAGAGAGGAAACAAAGGACCAAGG 1545
Db 597 CysLysGluLysGlyHisTrpAlaLysAspCysProLysLys---ArgArgGlnPheLys 615
QY 1546 ATCTAGCTCTAGAGAGAT 1566
Db 616 ValLeuThrLeuGluAspAsp 622

RESULT 15
AAB10043
ID AAB10043 standard; protein; 538 AA.
XX
AC AAB10043;
XX
DT 02-NOV-2000 (first entry)
XX
DE MMLV gag protein.
XX
KW Glycoprotein; gag gene; pol gene; GP-1; GP-2; anti-HIV; cytostatic;
KW gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;
KW carcinoma; melanoma.
XX
OS Moloney murine leukemia virus.
XX
FN EPI006196-A2.
XX
PD 07-JUN-2000.
XX
PF 25-NOV-1999; 99EP-00250415.
XX
PR 26-NOV-1998; 98DE-01056463.
XX
PA (PETT-) PETTE INST HEINRICH.
XX
PI Von Laer MD;
XX
DR WPI; 2000-378268/33.
XX
PT New retroviral packing cell useful as pharmaceutical carrier in gene
PT therapy for treatment of HIV and neoplasms, comprises retroviral genes
PT and glycoproteins.
```

```
XX
PS
XX Disclosure; Page 31-32; 69pp; German.
CC This invention describes a novel retroviral packing cell (I), comprising
CC the retroviral genes gag, pol and glycoproteins GP-1 and GP-2 of the LCMV
CC coding gene gp, or a part of these. The products of the invention have
CC anti-HIV and cytostatic activity and can be used for gene therapy. (I) is
CC useful for in vitro infection of cells, especially hematopoietic stem
CC cells, for expression of transgenes in cells and as a pharmaceutical
CC carrier for gene therapy. (I) is therefore useful in the treatment of
CC infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and
CC other diseases. This sequence represents the Moloney murine leukemia
CC virus (MMLV) gag protein which is described in the method of the
CC invention
XX
SQ Sequence 538 AA;
Alignment Scores:
Pred. No.: 1,96e-113 Length: 538
Score: 1380.50 Matches: 283
Percent Similarity: 62.3% Conservative: 69
Best Local Similarity: 50.0% Mismatches: 141
Query Match: 48.5% Indels: 73
DB: 3 Gaps: 13
US-10-723-552-3_COPY_585_2156 (1-1572) x AAB10043 (1-538)
QY 1 ATGGAGCAGAGCGGTGAGACCCCTCTAGTTTGTGATCTCGACCATTTGACTGAAGTTAAA 60
Db 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
QY 61 TCCAGGGCTCATATTTGTTCAGTTTCAGTTTAAAGAGGGACCTTCGCAGACTTCTGTCTC 120
Db 21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgArgTrpValThrPheCysSer 40
QY 121 TCTGAATGGCCGACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAAATTCAGAGATT 180
Db 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
QY 181 ATCTGGCTGTAAAGCAGATTATTTTCAGACTCGACCCGGCTCTCATCCGATCAGAG 240
Db 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
QY 241 CCTATATCTTACGTGGCAAGATTGGCAGAGATCTCGCCATCGGCTGGTGTAAACCATGG 300
Db 81 ProTyrIleValThrTrpGluAlaPheAspProProProTrpValLysProPhe 100
QY 301 CTGAATAAGCAAGAAAGCCAGGTCCTCCGAAATTCGTCTTGGAGAGAAACAAACAC 360
Db 101 ValHis---ProLysProProProProProProProSerAlaProSerLeuProLeuGlu 119
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGAGGCCA---CCG 417
Db 120 ProProArgSerThrProProArgSerSerLeuTyrProAlaLeuThrProSerLeuGly 139
QY 418 GCTTGGCCGGAACCCCA-----GGAGCTCCGGGTGGAGGACCT-----TCT 438
Db 140 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuIleAspLeuLeuThr 159
QY 439 GTTCCCCCAACCCCTTATCTGGCAGAGGTCGCGAGGGACCCCTTGGCCCTCTCT--- 495
Db 160 GluAspProProProTyr-----ArgAspProArgProProProSer 173
QY 496 -----GGAGCTCCGGGTGGAGGACCT----- 519
Db 174 AspArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyAlaProAspProSerPro 193
QY 520 ---GCTGCAGGACTCGAGCCGGAGGGGCCACCCCGAGCGGACACAGAGATCGCG 576
Db 194 MetAlaSerArgLeuArgGlyArgGluProProValAlaAspSerThrThrSerGln 213
QY 577 ACATTACCGCTGCGCAGCGTACGGCCCTCCACACCGGGGGGCCAATTTCAGCCCTCCAG 636
Db 577 ACATTACCGCTGCGCAGCGTACGGCCCTCCACACCGGGGGGCCAATTTCAGCCCTCCAG 636
```


Db 214 AlaPheProLeuArgAlaGlyGly-----AsnGlyGln-----LeuGln 226
QY 637 TATTGGCCCTTTCTTTCGAGATCTCTATAATTCGAAAACTAACCATCCCTTCTTCG 696
Db 227 TyrTrpProPheSerSerAspLeuTyrAsnTrpLysAsnAsnProSerPheSer 246
QY 697 GAGNATCCCAACGCTCACGGGGTGTGGAGTCCCTTATGTTCTCACAGCCTACT 756
Db 247 GluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuIleThrHisGlnProThr 266
QY 757 TGGGATGATTGTCAACAGCTGTCGAGACACTCTTTCACACCGGAGGCGAGAGAGATT 816
Db 267 TrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgVal 286
QY 817 CTATTAGAGCTAGAAAAAATGTTCTCGGGCCGACCGGGCGACCCACCGCGTTCGAAAAAT 876
Db 287 LeuLeuGluAlaArgLysAlaValArgGlyAspAspGlyArgProThrGlnLeuProAsn 306
QY 877 GAGATTGACATGGGATTTCCCTTAACCTGCCCGGTTGGGACTACACACCGCTGAAGGT 936
Db 307 GluValAspAlaAlaPheProLeuGluArgProAspTrpAspTyrThrGlnAlaGly 326
QY 937 AGGAGAGCTTGAAAAATCTATCGCCAGGCTCTGTGGCGGTCTCCGGGGCGCTCAAGA 996
Db 327 ArgAsnHisLeuValHisTyrArgGlnLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 346
QY 997 CGGCCCACTAATTTGGCTAAGGTAAAGAGTGAATGCAGGAGCGCAATGAACCCCTCT 1056
Db 347 SerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer 366
QY 1057 GTTTTCTTGAGAGCTCTTGAAGCCTTCAGCGGTACACCCCTTTGATCCCACTCA 1116
Db 367 AlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp 386
QY 1117 GAGGCCCAAAAGCCTCAGTCGCTTGGCCTTTATAGACAGTCAGCCTTGATATTAGA 1176
Db 387 ProGlyGlnGluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGly 406
QY 1177 AAGAAGCTTCAGAGACTGGAAGGTTACAGAGGCTGAGTTACGTGATCTAGTGAAGGAG 1236
Db 407 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 426
QY 1237 GCAGAGAAATATTATCAAAAGGAGACAGAGAAAGGAAACAAAGAAAGAGAGA 1296
Db 427 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgIleArgArg 446
QY 1297 GAAGAGAGGAAGGAGGAAGA-----CGT 1323
Db 447 GluThrGluGluLysGluGluArgArgThrGluAspGluGlnLysGluLysGluArg 466
QY 1324 AATAACGGCAAGAGAAATTTGACTAAGATCTTGGCTGAGTGGTGAAGGAAAGC 1383
Db 467 AspArgArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys 486
QY 1384 AATACGGAAGAGAGAGAGATTTTAGGAAATTAGTCAGGCCCTAGACAGTCAGGGAAC 1443
Db 487 Gln-----AspArgGlnGlyGlu 493
QY 1444 CTGGGCAATAGACCCCACTCGACAGGACCAATGTCATATTGTAAGAGAGGACAC 1503
Db 494 ---ArgArgArgSerGlnLeuAspArgGlnCysAlaTyrCysLysGluLysGlyHis 512
QY 1504 TGGCAAGAGAACTCCCAAGAAG-----GGAACAAAGGACCAAG----- 1545
Db 513 TrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArgProGlnThrSer 532
QY 1546 ATCTAGCTCTAGAAGAA 1563
Db 533 LeuLeuThrLeuAspAsp 538

This Page Blank (uspto)

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:07:51 ; Search time 39.5271 Seconds
(without alignments)
3323.427 Million cell updates/sec

Title: US-10-723-552-3_COPY_585_2156

Perfect score: 2848

Sequence: 1 ATGGGACAGACGGTGACGAC.....CTCTAGAGAGATTAAGAT 1572

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10723552/runat.14022006.125148.13205/app.query.fasta.1
-DB=Published Applications AA Main -QFMT=fascan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN 1.1 602 @runat.14022006.125148.13205 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380.5	48.5	538	4	US-10-677-558-1
2	1379.5	48.4	538	3	US-09-827-822-9
3	967	34.0	957	5	US-10-732-923-13461
4	832.5	29.2	253	4	US-10-670-695-28
5	825	29.0	774	5	US-10-732-923-13451
6	573	20.1	763	5	US-10-732-923-10649
7	553	19.4	405	5	US-10-450-763-38327
8	553	19.4	430	5	US-10-450-763-39297
9	553	19.4	1577	5	US-10-450-763-36834
10	553	19.4	1577	5	US-10-450-763-57482
11	543.5	19.1	611	4	US-10-670-695-32
					Sequence 1, Appli
					Sequence 9, Appli
					Sequence 13461, A
					Sequence 28, Appl
					Sequence 13451, A
					Sequence 10649, A
					Sequence 38327, A
					Sequence 39297, A
					Sequence 36834, A
					Sequence 57482, A
					Sequence 32, Appl

12	542.5	19.0	878	5	US-10-450-763-36843	Sequence 36843, A
13	526.5	18.5	688	5	US-10-450-763-39289	Sequence 39289, A
14	497	17.5	918	5	US-10-732-923-13456	Sequence 13456, A
15	497	17.5	981	5	US-10-732-923-13455	Sequence 13455, A
16	478	16.8	505	5	US-10-450-763-45672	Sequence 45672, A
17	475.5	16.7	208	4	US-10-029-386-33154	Sequence 33154, A
18	439.5	15.4	368	5	US-10-450-763-39293	Sequence 39293, A
19	439.5	15.4	606	5	US-10-450-763-39335	Sequence 39335, A
20	439.5	15.4	606	5	US-10-450-763-39288	Sequence 39288, A
21	439.5	15.4	652	5	US-10-450-763-38335	Sequence 38335, A
22	435.5	15.3	510	5	US-10-450-763-38333	Sequence 38333, A
23	415.5	14.6	202	4	US-10-029-386-32299	Sequence 32299, A
24	405.5	14.1	793	5	US-10-450-763-57559	Sequence 57559, A
25	397.5	14.0	402	5	US-10-450-763-36213	Sequence 36213, A
26	392.5	13.8	997	5	US-10-450-763-39301	Sequence 39301, A
27	392	13.8	1015	5	US-10-450-763-34486	Sequence 34486, A
28	387.5	13.6	1209	5	US-10-450-763-36212	Sequence 36212, A
29	386	13.6	693	5	US-10-450-763-39299	Sequence 39299, A
30	369	13.0	309	5	US-10-450-763-45680	Sequence 45680, A
31	359.5	12.6	245	3	US-09-864-761-36722	Sequence 36722, A
32	355.5	12.5	620	5	US-10-450-763-33954	Sequence 33954, A
33	355.5	12.5	620	5	US-10-450-763-36206	Sequence 36206, A
34	355.5	12.5	620	5	US-10-450-763-36845	Sequence 36845, A
35	355.5	12.5	620	5	US-10-450-763-45682	Sequence 45682, A
36	352.5	12.4	809	5	US-10-450-763-36849	Sequence 36849, A
37	348	12.2	1035	5	US-10-450-763-35678	Sequence 35678, A
38	328	11.5	486	5	US-10-450-763-36846	Sequence 36846, A
39	328	11.5	486	5	US-10-450-763-52359	Sequence 52359, A
40	313.5	11.0	420	4	US-10-637-565-15	Sequence 15, Appl
41	310	10.9	544	5	US-10-450-763-35688	Sequence 35688, A
42	310	10.9	544	5	US-10-450-763-47717	Sequence 47717, A
43	310	10.9	550	5	US-10-450-763-40946	Sequence 40946, A
44	309	10.8	507	5	US-10-450-763-52789	Sequence 52789, A
45	302.5	10.6	390	5	US-10-450-763-39975	Sequence 39975, A

ALIGNMENTS

RESULT 1

US-10-677-558-1
; Sequence 1, Application US/10677558
; Publication No. US20040096972A1
; GENERAL INFORMATION:
; APPLICANT: AUDIT, Muriel
; APPLICANT: COSSET, Francois-Loic
; TITLE OF INVENTION: CHIMERIC PLASMID COMPRISING A REPLICATIVE RETROVIRAL GENOME AND
; FILE REFERENCE: 1759.135
; CURRENT APPLICATION NUMBER: US/10/677,558
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/FR02/03934
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: FR 0114976
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: product of gag sequence of pAM plasmid
US-10-677-558-1

Alignment Scores:
Pred. No.: 6.8e-95 Length: 538
Score: 1380.50 Matches: 283
Percent Similarity: 62.2% Conservative: 69
Best Local Similarity: 50.0% Mismatches: 141
Query Match: 48.5% Indels: 73
DB: 4 Gaps: 13

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-677-558-1 (1-538)

```
QY 1 ATGGGACAGCGGTGACGACCCCTCTAGTTGTTGACTCTCGACCATTTGAGCTGAAGTTAAA 60
Db |||||
QY 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGlyHisTrpLysAspValGlu 20
Db |||||
QY 61 TCCAGGGCTCATAAATTCTCAGTTTCAAGTAAAGGACCTTTGGCAGACTTTCTGTGTC 120
Db |||||
QY 21 ArgIleAlaHisAsnGlnSerValAspValLysLysArgTTrpValThrPheCysSer 40
Db |||||
QY 121 TCTGAATGCCGACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAAATCTCAGATT 180
Db |||||
QY 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
Db |||||
QY 181 ATCTGGCTGTAAAGCAGCTATTATTTTCAGACTGGACCCGGCTCTCATCCCGATCAGGAG 240
Db |||||
QY 61 IleThrGlnValLysIleLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
Db |||||
QY 241 CCTATATCTTACGTGTCAGATTTGCGAGGATCTCCGCCATGGTTTAAACCATGG 300
Db |||||
QY 81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProTrpValLysProPhe 100
Db |||||
QY 301 CTGAATAGCCAGAAAGCCAGGTCCTCCGATTTCTGGCTCTTGGAGAAACAAACAC 360
Db |||||
QY 101 ValHis---ProLysProProProLeuProProSerAlaProSerLeuProLeuGlu 119
Db |||||
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCGAGATTTGAGGAGCCA---CCG 417
Db |||||
QY 120 ProProArgSerThrProProArgSerSerLeuThrProAlaLeuThrProSerLeuGly 139
Db |||||
QY 418 GCTTGGCCGGAACCCCAA-----GGAGCTCCGGGGTGGAGGACCT-----TCT 438
Db |||||
QY 140 AlaLysProLysProGlnValLeuSerAspSerGlyGlyProLeuIleAspLeuLeuThr 159
Db |||||
QY 439 GTTCCCCCACCCTTATCTGCGACAGGTCGGCGGAGGACCTTTGCCCTCCT--- 495
Db |||||
QY 160 GluAspProProProTyr-----ArgAspProArgProProSer 173
Db |||||
QY 496 -----GGAGCTCCGGGGTGGAGGACCT----- 519
Db |||||
QY 174 AspArgAspGlyAsnGlyGlyGluAlaThrProAlaGlyGluAlaProAspProSerPro 193
Db |||||
QY 520 ---GTCGAGGAGACTCGGAGCGCGGAGCGCCACCCCGGAGCGGACAGACGAGATCCG 576
Db |||||
QY 194 MetAlaSerArgLeuArgGlyArgArgGluProProValAlaAspSerThrSerGln 213
Db |||||
QY 577 ACATTACCGCTCGGACGACTACGGCCCTCCACACCGGGGGGCAATTCACGCCCTCCAG 636
Db |||||
QY 214 AlaPheProLeuArgAlaGlyGly-----AsnGlyGln-----LeuGln 226
Db |||||
QY 637 TATTGGCCCTTTCTCTCGACATCTCTATAATTGGAATACTAACCATCCCTCTTCG 696
Db |||||
QY 227 TyrTrpProPheSerSerAspLeuTyrAsnTrpLysAsnAsnProSerPheSer 246
Db |||||
QY 697 GAGGATCCCAACCGCTCACCGGGTGTGGAGTCCCTTATGTTCTCTCACCAGCTACT 756
Db |||||
QY 247 GluAspProGlyLysLeuThrAlaLeuIleGluSerValLeuThrHisGlnProThr 266
Db |||||
QY 757 TGGGATGATTGTCAACAGCTGTCGACACTCTTCACACCGGAGCGGAGAGAGATT 816
Db |||||
QY 267 TrpAspAspCysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgVal 286
Db |||||
QY 817 CTATTAGAGGCTAGAAAAAATGTTCTCGGGCCGAGCGGACCCACCGGCTTGCATAAT 876
Db |||||
QY 287 LeuLeuGluAlaArgLysAlaValArgGlyAspAspGlyArgProThrGlnLeuProAsn 306
Db |||||
QY 877 GAGATTGACATGGGATTTCCCTTAACTCGCCCCGTTGGGACTACACACCGGCTGAAGGT 936
Db |||||
QY 307 GluValAspAlaAlaPheProLeuGluArgProAspTrpAspTyrThrThrGlnAlaGly 326
Db |||||
QY 937 AGGAGAGACTTGAATAATCTATCGCAGGCTCTGTTGGGGGTCTCCGGGGGCGCTCAGA 996
Db |||||
QY 327 ArgAsnHisLeuValHisTyrArgGlnLeuLeuLeuAlaGlyLeuGlnAsnAlaGlyArg 346
Db |||||
QY 997 CGGCCCCACTAATTTGGCTAAGGTAAAGAGAGATGATGCAGGGACCGAATGAACCCCTCT 1056
Db |||||
```

```
Db |||||
QY 347 SerProThrAsnLeuAlaLysValLysGlyIleThrGlnGlyProAsnGluSerProSer 366
Db |||||
QY 1057 GTTTTCTTGAGAGCTCTTGAAGCCTTACAGCGGTACACCCCTTTTGATCCCACTCA 1116
Db |||||
QY 367 AlaPheLeuGluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAsp 386
Db |||||
QY 1117 GAGGCCCAAAAGCCTCAGTGGCTTGGCTTTATAGGACAGTCCAGCTTCGATATTAGA 1176
Db |||||
QY 387 ProGlyGlnGluThrAsnValSerWetSerPheIleTrpGlnSerAlaProAspIleGly 406
Db |||||
QY 1177 AAGAGCTTTCAGAGACTGGAAGGTTTACAGAGAGCTGAGTTTACGTGATCTAGTGAAGGAG 1236
Db |||||
QY 407 ArgLysLeuGluArgLeuGluAspLeuLysAsnLysThrLeuGlyAspLeuValArgGlu 426
Db |||||
QY 1237 GCAGAGAAAGTATTATTACAAAAGGAGACAGAAAGAAAGGAAACAAAGAAAGAGAGA 1296
Db |||||
QY 427 AlaGluLysIlePheAsnLysArgGluThrProGluGluArgGluGluArgIleArgArg 446
Db |||||
QY 1297 GAAAGAGAGGAAAGGAGGAAAGA-----CGT 1323
Db |||||
QY 447 GluThrGluGluLysGluArgArgThrGluAspGluGlnLysGluLysGluArg 466
Db |||||
QY 1324 AATAAAGCGCAAGAGAAGAAATTTGACTAAGATCTTGGCTGCAGTGGTTGAAGGGAAGAC 1383
Db |||||
QY 467 AspArgArgArgHisArgGluMetSerLysLeuLeuAlaThrValValSerGlyGlnLys 486
Db |||||
QY 1384 AATACGGAAGAGAGAGAGATTTAGGAAAATTAGTCCAGGCCCTTAGACAGTCAGGGAAC 1443
Db |||||
QY 487 Gln-----AspArgGlnGlyGlyGlu 493
Db |||||
QY 1444 CTGGGCAATAGGACCCCACTCGACAAGGACCAATGTGCATATTGTAAAGAAAGAGAGACAC 1503
Db |||||
QY 494 ---ArgArgArgSerGlnLeuAspArgAspGlnCysAlaTyrCysLysGluLysGlyHis 512
Db |||||
QY 1504 TGGGCAAGAACTGCCCCCAAGAAG-----GGAACCAAGGACCAAGG----- 1545
Db |||||
QY 513 TrpAlaLysAspCysProLysLysProArgGlyProArgGlyProArgProGlnThrSer 532
Db |||||
QY 1546 ATCTAGCTCTAGAAGAA 1563
Db |||||
QY 533 LeuLeuThrLeuAspAsp 538
Db |||||

RESULT 2
US-09-827-822-9
; Sequence 9, Application US/09827822
; Patent No. US20020091086A1
; GENERAL INFORMATION:
; APPLICANT: Anders Vahline
; TITLE OF INVENTION: PEPTIDES THAT BLOCK VIRAL INFECTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.003A
; CURRENT APPLICATION NUMBER: US/09/827,822
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US/09/370,368
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Moloney Murine Leukemia Virus
US-09-827-822-9

Alignment Scores:
Pred. No.: 8,09e-95 Length: 538
Score: 1379.50 Matches: 280
Percent Similarity: 61.8% Conservative: 77
Best Local Similarity: 48.4% Mismatches: 124
Query Match: 48.4% Indels: 97
DB: 3 Gaps: 13

US-10-723-552-3_COPY_585_2156 (1-1572) x US-09-827-822-9 (1-538)
```


US-10-670-695-28

Alignment Scores:

Pred. No.: 8,36e-54 Length: 253
Score: 832.50 Matches: 159
Percent Similarity: 77.6% Conservative: 38
Best Local Similarity: 62.6% Mismatches: 50
Query Match: 29.2% Indels: 7
DB: 4 Gaps: 2

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-670-695-28 (1-253)

```
QY 580 TTACGGCTCGCAGCTACGGCCCTCCACACGGGGGCCCAATTGCGACCCCTCAGTAT 639
DB 1 MetProLeuArg-----GluGlyProAsnAsnArgPro---GlnTyr 13

QY 640 TGGCCCTTTCTTCGACAGATCTCTATAATTGAAACTAACCATCCCTTTCTCCGAG 699
DB 14 TrpProPheSerAlaSerAspLeuTyrAsnTrpLysSerHisAsnProProPheSerGln 33

QY 700 GATCCCCAACCCCTCACGGGGTGGTGGAGTCCCTTATGTCTCTCACAGCCTACTTGG 759
DB 34 AspProValAlaLeuThrAsnLeuIleGluSerIleLeuValThrHisGlnProThrTyr 53

QY 760 GATGATGTTCAACAGCTCTCGACACATCTTCAACCCGAGGAGCGAGAGAAATCTTA 819
DB 54 AspAspCysGlnGlnLeuLeuGlnAlaLeuLeuThrGlyGluGluArgGlnArgValLeu 73

QY 820 TTAGAGGCTAGAAAAATGTTCTTGGGCCCGACGGGACCCACCGGTTGCAAAATGAG 879
DB 74 LeuGluAlaArgLysGlnValProGlyGluAspGlyArgProThrGlnLeuProAsnVal 93

QY 880 ATTGACATGGGATTTCCTTAACTCGCCCGGTTGGGACTACACACGGCTGAAGGTAGG 939
DB 94 IleAspGluAlaPheProLeuThrArgProAsnTrpAspPheAlaThrProAlaGlyArg 113

QY 940 GAGAGCTTGAAATCTATCGCCAGCTCTGTGGGGGTCTCCGGGGGCTCAAGAGGG 999
DB 114 GluHisLeuArgLeuTyrArgGlnLeuLeuAlaGlyLeuArgGlyAlaAlaArgArg 133

QY 1000 CCCACTAATTGGCTTAAGTAAGAGTATGATGACGAGGACCGAATGACCCCTCTGTT 1059
DB 134 ProThrAsnLeuAlaGlnValLysGlnValGlnGlyLysGluGluThrProAlaSer 153

QY 1060 TTTCTTGAGAGGCTCTTGGAGCCCTCAGGGCGGTACACCCCTTTTGATCCACCTCAG 1119
DB 154 PheLeuGluArgLeuLysGluAlaTyrArgMetTyrThrProTyrAspProGluAspPro 173

QY 1120 GCCCAAAAGCCTCAGTGGCTTTGGCCCTTTATAGGACAGTCAGCTTGGATATPAGAA 1179
DB 174 GlyGlnAlaAlaSerValIleLeuSerPheIleTyrGlnSerSerProAspIleArgAsn 193

QY 1180 AAGCTTCAGAGCTGGAAGGTTACAGAGGCTGAGTTCGTGATCTAGTGAAGGAGCA 1239
DB 194 LysLeuGlnArgLeuGluGlyLeuGlnGlyPheThrLeuSerAspLeuLeuLysGluAla 213

QY 1240 GAGAAAGTATATTACAAAGGAGACAGAGAGAAAGGAGAAAGGAGAGAGAGAA 1299
DB 214 GluLysIleTyrAsnLysArgGluThrProGluGluArgGluGluArgLeuTyrGlnArg 233

QY 1300 AGAGAGAAAGGAGAGAAAGCTGTAATAAAGCGCAAGAGAG 1341
DB 234 GlnGluGluArgAspLysArgHisLysGluMetThrLys 247
```

RESULT 5

US-10-732-923-13451
; Sequence 13451, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923

; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 13451
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Feline sarcoma virus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(774)
; OTHER INFORMATION: unsure at all Xaa locations
; US-10-732-923-13451

Alignment Scores:
Pred. No.: 3.93e-53 Length: 774
Score: 825.00 Matches: 191
Percent Similarity: 53.1% Conservative: 50
Best Local Similarity: 42.1% Mismatches: 135
Query Match: 29.0% Indels: 78
DB: 5 Gaps: 14

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-732-923-13451 (1-774)

```
QY 1 ATGGGACAGACGGTACGACCCCTTTAGTTTGAATCTCGACCAATTGGACTGAAGTTAAA 60
DB 1 MetGlyGlnThrValThrProLeuSerLeuThrLeuAspHisTrpSerGluValArg 20

QY 61 TCCAGGCTCATATTTGTCAAGTTAGTAAAGAGGACCTTGGCAGACTTCTGTGTC 120
DB 21 AlaArgAlaHisAsnGlnGlyValGluValArgLysLysIleThrLeuCysLys 40

QY 121 TCTGAATGGCGCACATTCGATGTTGGATGGCCATCAGAGGGACCTTTAATTCAGATT 180
DB 41 AlaGluTrpValMetMetAsnValGlyTrpProArgGluGlyThrPheSerLeuAspAsn 60

QY 181 ATCTGCTGCTTAAAGCAGTTATTTTCAGATGGACCCGCTCTCATCCGATCAGAG 240
DB 61 IleSerGlnValLysLysIlePheAlaProGlyProHisGlyHisProAspGlnVal 80

QY 241 CCCTATATCTTACGTGGCAGATTGGCAGAGGATCTCCGCCATCGGTAAACATGG 300
DB 81 ProTyrIleThrThrTrpArgSerLeuAlaThrAspProSerTrpValArgProPhe 100

QY 301 CTGAATAAGCAAGAACCCAGGTCCCGGAATTCGTGCTCTTGAGAGAGAAAAACAACAC 360
DB 101 LeuProProLysProProThrPro----- 109

QY 361 TCGGCTGAAAAAGTCAAGCCCTCTCTCT-----CATATCTACCCGAG 402
DB 110 LeuProGlnProLeuSerProGlnProSerAlaProLeuThrSerSerLeuTyrProVal 129

QY 403 ATT-----GAGGAGCCACCGCTTGGCCGGAACCCCAATCTGTTCCC--- 444
DB 130 ValProLysProAspProProLysProProValLeuProProAspProSerSerProLeu 149

QY 445 -----CACACCCCTTATCTGSCACAGGTGCGCGAGGGAGCCC 483
DB 150 IleAspLeuLeuThrGluGluProProProTyr-----ProGlyGlyHisGlyProPro 167

QY 484 TTTGCCCCCTCTCGAGCTCCGGCGGTGGAGGACCTGTCTCAGGGACTCGGAGCCGAGG 543
DB 168 ProSerGlyProArgThrProAlaAlaSerProIleValSerArgLeuArgGluArg 187

QY 544 GGGCCACCCCGGAGCGGACAGAGATCGGACATTACCGTTCGCCACGATACGGCCCT 603
DB 188 -----GluAsnProAlaGluGluSerGlnAlaLeuProLeuArg----- 200

QY 604 CCCACACCGGGGGCCAAATTGACGCCCTCCAGTATTGGCCCTTTCTTCTCGAGATCTC 663
DB 201 -----GluGlyProAsnAsnArgPro---GlnTyrTrpProPheSerAlaSerPhe 217

QY 664 TATAATTGGAATAACTAACCATCCCTTTCTCGGAGGATCCCAACAGCCTCAGGGGTTG 723
```

```
Db      218  TyrAsnTrpLysSerHisAsnProProPheSerGlnAspProValAlaLeuThrAsnLeu 237
      724  GTGGAGTCCCTTATGTTCTCCACAGCCTACTTGGGATGATGTCACAGCTCTGCGAG 783
      238  ileGluSerileLeuValThrHisGlnProThrTrpAspAspCysGlnGlnLeuGln 257
      784  ACACCTCTTCACACCGGAGGAGGAGAGAGATTTCTATTAGAGGCTAGAAAAATGTTCT 843
      258  AlaLeuLeuThrGlyGluGluArgValLeuLeuGluAlaArgLysGlnValPro 277
      844  GGGCGCAGCGCGCACCCACCGGTTGCCAAATGAGATTTGACATGGGATTTCCCTTAAC 903
      278  GlyGluAspGlyArgProThrGlnLeuProAsnValIleAspGluThrPheProLeuThr 297
      904  CGCCCGGTTGGGACTACAAACCGCTGAAGGTAGGGAGACTTGAAATCTATCGCCAG 963
      298  HisProArgGluGlnValGlnLeuLeuAlaLysLysGlnValLeuGlnAlaLeuGln 317
      964  GCTCTGGTGGGGTCTCCGGGGCGCTCAAGACGGCCACTAATTGGCTTAAGGTAGA 1023
      318  AlaLeuGlnValAlaLeuCysSerGlnAlaLys-----LeuGlnAlaGlnArg 333
      1024  GAAGTGATGCAG-----GGACCGAATGAACCCCTCTGTTTCTT 1065
      334  GluLeuLeuGlnAlaLysLeuGluGlnLeuGlyProGlyGluProProValLeu--- 352
      1066  GAGAGGCTCTTGGAGCCCTCAGCGGTACACCCCTTTTGAT----- 1107
      353  -----LeuLeuGlnAspAspArgHisSerThrSerSerSerGluGlnGluGly 370
      1108  -----CCCACTCAGAGCCCAAAAGCCTCAGTGGCTTTGGCCTTTATAGGACAG 1158
      371  GlyArgThrProThrLeuGluIleLeuLysSerHisIleSerGlyIlePhe----- 387
      1159  TCAGCCTTGGATATTAGAAAGAGCTTCAGAGACTGGAAGGTTACAGGAGGCTGAGTTA 1218
      388  -----ArgProLysPheSerLeuProProProLeuGln----- 398
      1219  CGTGATCTAGTAGGAGGCGAGAGAAAGTATATTACAAAGG 1260
      399  -----LeuValProGluValGlnLysProLeuHisGluGln 410
```

RESULT 6

```
US-10-732-923-10649
; Sequence 10649, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10649
; LENGTH: 763
; TYPE: PRT
; ORGANISM: AKT8 murine leukemia virus
US-10-732-923-10649
```

```
Alignment Scores:
Pred. No.: 3,33e-34 Length: 763
Score: 573.00 Matches: 174
Percent Similarity: 42.2% Conservative: 68
Best Local Similarity: 30.4% Mismatches: 146
Query Match: 20.1% Indels: 185
DB: 5 Gaps: 25
```

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-732-923-10649 (1-763)

```
Db      1  MetGlyGlnThrValThrThrProLeuSerLeuThrLeuGluHisTrpGlyAspValGln 20
      61  TCCAGGGCTCATATTTGTTCAGTTTCAAGAGGACCTTGGCAGACTTTCTGTGTC 120
      21  ArgIleAlaSerAsnGlnSerValAspValLysLysArgArgTrpValThrPheCysSer 40
      121  TCTGAATGGCGGACATTCGATGTGGATGGCCATCAGAGGGGACCTTTAATCTCGAGATT 180
      41  AlaGluTrpProThrPheGlyValGlyTrpProGlnAspGlyThrPheAsnLeuAspIle 60
      181  ATCTCTGGCTGTAAAGCAGATTATTTTTCAGACTGGACCCCGCTCTCATCCCGATCAGGAG 240
      61  IleLeuGlnValLysSerLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
      241  CCTATATCTCTTACGTGGCAAGATTGGCAGAGGATCTTCGCCCATGGTGTAAACCATGG 300
      81  ProTyrIleValThrTrpGluAlaIleAlaLysArgLysProProTrpValLysProPhe 100
      301  CTGAATAAGCCNAGAAAGCCAGGTCCCGAATTCTGGCTCTTGGAGAGAAAAACAACAC 360
      101  ValSerProLysLeuSerLeuSerProThrAlaProIleLeu-----Pro 115
      361  TCGGCTGAAAAAGTCAAGCCCTCTCTCAT-----ATCTACCCCGAG 402
      116  SerGlyProSerThrGlnProProProArgSerAlaLeuTyrProAlaLeuThrProSer 135
      403  ATTGAGGAGCCACCGGCTTGGCGGAACCCCAATCT----- 438
      136  IleLysProArgProSerLysProGlnValLeuSerAspAsnGlyGlyProLeuIleAsp 155
      439  -----GTTCCCCACACCCCTTATCTGSCACAGGCT----- 468
      156  LeuLeuThrGluAspProProProTyrGlyGluGlnGlyProSerSerSerAspGlyAsp 175
      469  -----GCCGCGAGGGGACCTTTCCTCCCTCTCTGGAGCTCCCGCGGTG--- 510
      176  GlyAspArgGluGluAlaThrSerThrProGluIleProAlaProSerProMetValSer 195
      511  -----GAGGGACCTGCTGACGGGACTCGGAGCCGAGCGGCGGCCACCC 552
      196  ArgLeuArgGlyLysArgAspProProAlaAlaValSerThrThrSerArg----- 212
      553  CCGGAGCGGACAGACGAGATCGGCACATTACCGCTGCGCAGCTACGCGCCCTCCACACCG 612
      213  -----AlaPheProLeuArgLeuGlyGly----- 220
      613  GGGGGCCAAATGACGCCCTCCAGTATTGGCCCTTTTCTCTGCAGATCTCTATAATTGG 672
      221  AsnGlyGln-----LeuGlnTyrTrpProPheSerSerSerAspLeuTyrAsnTrp 237
      673  AAAACTAACCATCCCTTCTCGGAGATCCCAACGCCTCACGGGTTGGTGGAGTCC 732
      238  LysAsnAsnAsnProSerPheSerGluAspProGlyLysLeuThrAlaLeuIleGluSer 257
      733  CTTATGTTCTCTACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTC 792
      258  ValLeu----- 259
      793  ACAACCGGAGGAGCGAGAGAGAAATTTCTATTAGAGGCTAGAAAAAATGTTCTGGGGCCGAC 852
      260  ThrThrHisAlaArgGluGluThrLeuIle-----IleIleProGlyLeu--- 274
      853  GGGCGACCC-----ACGGGTTGCAAAATGAGATTGACATGGGATTTCCC 897
      275  -----ProLeuSerLeuGlyAlaThrAspThrMetAsnAspVal-----Ala 288
      898  TTAACCTGCCCGCTGGGACTACAAACGCGCTGAAGGTAGGAGAGCTTGAAATCTAT 957
      289  IleValLysGluGlyTrpLeuHisLysArgGly-----GluTyrIleLysThrTrp 305
      958  CGCCAGGCTCTGTGTG-----GCCGGTCTCCGGGGCGCTCAAGACGGCCC 1002
```

QY 1 ATGGGACAGACGGTGCAGCACCCCTCTTAGTTTGACTCTCGACCATTTGGACTGAAGTTAAA 60

Db 306 ArgProArgTyrPheLeuLeuLysAsnAspGlyThrPheIleGlyTyrLysGluArgPro 325
Qy 1003 ACTAATTGGCTAAGTAAGA-----GAAGTG 1029
Db 326 GlnAspValAspGlnArgGluSerProLeuAsnAsnPheSerValAlaGlnCysGlnLeu 345
Qy 1030 ATGAGGAGACCAAGTAAGACCCCTCTCTTTCTTGAGAGGCTCTGGAA----- 1080
Db 346 MetLysThrGluArgProArgProAsnThrPheIleIleArgCysLeuGlnTrpThr 365
Qy 1081 -----GCCTTCAGCGGTATACACCCCTTTGATCCACCTCAGAGGCCCAAAAA 1128
Db 366 VallIleGluArgThrPheHisValGluThrPro-----GluGluArgGluGln 381
Qy 1129 GCCTCAGTGGCTTTGGCCTTTATAGGACAGTCAGCCTTTGGATATTAGAAAGAGCTTCAG 1188
Db 382 TrpAlaThrAlaIleGlnThrValAla----- 390
Qy 1189 AGACTGGAGGTTTACAGGAGCTGAGTTACGTGATCTAGTGAAGGAGCGACAGAAAGTA 1248
Db 391 -----AspGlyLeu----- 393
Qy 1249 TATTACAAAGGGAGACAGAGAAGAAAGGGAACAAAGAAAA-----GAG 1293
Db 394 -----LysArgGlnGluGluThrMetAspPheArgSerGlySerProSerAspAsn 411
Qy 1294 AGAAAGAGAGAAAGGAGGAGAGACGTAATAACCGCAAGAGAAAGAAATTTGACTAAG 1353
Db 412 SerGlyAlaGluGluMetGlu-ValSerLeuAlaLysProLysHisArgValThrMetAs 431
Qy 1354 ATCTTGCTGCTGCTGTTGAAGGGAAGAAACAAATACGGAAGAGAGAGAGATTTAGGAAA 1413
Db 431 nGluPhe---GluTyrLeuLys-----LeuLeuGlyLy 441
Qy 1414 ATTAGTCCAGGCCCTAGACAGTCAGGAGCACTGGGCAATAGGACCCCACTCGACAAGAC 1473
Db 441 s-----GlyThrPheGly----- 445
Qy 1474 CAATGTGCATATTGAAAGAAAGAGACACTGGGCAA 1510
Db 446 -LysValIleLeuValLysGluLysAlaThrGlyArg 457
RESULT 7
US-10-450-763-38327
; Sequence 38327, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 38327
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (106)..(311)
; OTHER INFORMATION: Gag p30 core shell protein domain identified by PFam,
; OTHER INFORMATION: accession name Gag_p30, E-value=1.2e-67, PFam score of 238.2
US-10-450-763-38327
Alignment Scores:
Pred. No.: 9.18e-33 Length: 405

Score: 553.00 Matches: 151
Percent Similarity: 48.6% Conservative: 64
Best Local Similarity: 34.2% Mismatches: 149
Query Match: 19.4% Indels: 78
DB: 5 Gaps: 19
US-10-723-552-3_copy_585_2156 (1-1572) x US-10-450-763-38327 (1-405)
Qy 298 TGGCTGAATTAAGCCCAAGAAAGCCAGGTCCTCCGAATTCCTGGCTCTTGGAGAGAAAAACAAA 357
Db 14 TrpLeuAsnThrGlyLysGlu-----LeuAlaLeuGlyValTrpThrPhe 28
Qy 358 CACTCGCGTGAAGAAAGTCAAGCCC-----TCTCCT-----CATATCTACCCCGAG 402
Db 29 GluThrTrpGluArgGlySerProLeuLeuSerProArgGlyLeuHisLeuHis----- 46
Qy 403 ATTGAGGAGCCACCGGCTTGGCGGAACCCCAATCTGTCTCCCCACCCCTTATCTGCGCA 462
Db 47 -----LysThrAsnIleProArgProProArgValGly 57
Qy 463 CAG-----GCTGCCGCGGAGGAGCCCTTTGGCCCTCTCTGGAGCTCCGCGGTGGAGGA 516
Db 58 LysArgGlyGlyGluAlaSerGlyGlu---ThrProPro----- 69
Qy 517 CCTGCTCAGGACTCGAGCCGAGGGGCGCCACCCCGAGCGGACAGAGATCGCG 576
Db 70 LeuAlaAlaGlyLeuArgProLysThrGly-----Ile 80
Qy 577 ACATTACCGTGGCAGC-----TACGGCCCTCCACACCGGGGGGCCAATTG----- 624
Db 81 GlnMetProLeuArgGluGlnTrpTyrThrGlyIleAspGluAspGlyHisMetValGlu 100
Qy 625 ---CAGCCCTCCAGTATTGGCCCTTTCTTCGCAGATCTCTATAATTTGAAAACTAAC 681
Db 101 ArgArgValPheValTyrGlnProPheSerAlaAspLeuAsnTrpLysAsnAsn 120
Qy 682 CATCCCTTTCTCGGAGGATCCCAACGCTCAACGGGTTGGTGGAGTCCCTTATGTC 741
Db 121 ThrProSerTyrThrGluLysProGlnAlaLeuIleAspLeuLeuGlnThrIleIleGln 140
Qy 742 TCTCACCAGCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCACACCCGAG 801
Db 141 ThrHisAsnProThrTrpAlaAspCysHisGlnLeuLeuMetPheLeuPheAsnThrAsp 160
Qy 802 GAGCGAGAGAGAAATCTATTAGAGGCTAGAAAA-----AATGTCCTGGGGCC 849
Db 161 GluArgArgValLeuGlnAlaAlaThrLysTrpLeuGluGluHisAlaProAlaAsp 180
Qy 850 GACGGCGACCCACGCGGTTGCAAAATGAGATTGACATGGGATTCCTCTTAACCTCGCCCC 909
Db 181 TyrGlnAsnPro-----GlnGluTyr-ValArgThrGlnLeuProGlyThrAspPro 197
Qy 910 GGTGGGACTACAACACGCGCTGAAGGTAGGAGAGCTTGAAATCTATCGCAGGCTCTG 969
Db 198 GlnTrpAspProAsnGluArgGluAspMetGlnArgLeuAsnArgTyrArgGluAlaLeu 217
Qy 970 GTGCGCGGCTCTCGGGCGCGCTCAAGACGCCCACTAAATTTGGCTAAGTAAGAGAGTG 1029
Db 218 LeuGluGlyLeuLysArgGlyAlaGlnLysAlaThrAsnValAla----- 232
Qy 1030 ATCAGGAGCGCAATGAACCCCTCTGTTTTTTTGTGAGAGGCTCTTTGGAAGCTTCAGG 1089
Db 233 IleGlnGlyLysGluGluSerProAlaGlnPheTyrGluArgLeuCysGluAlaTyrCys 252
Qy 1090 CGGTACACCCCTTTGATCCCACTCAGAGGCCCAAAAGCCCTCAGTGGCTTTGGGCTTT 1149
Db 253 MetTyrThrProPheAspProAsnSerProGluAsnGlnArgMetIleAsnMetAlaLeu 272
Qy 1150 ATAGGACAGTCAGCCTTGGATATTAGAAAGACTTCAGAGACTGGNAGGGTTACAGGAG 1209
Db 273 ValSerGlnSerAlaGluAspIleArgArgLysGlnAlaGlyPheAlaGly 292
Qy 1210 GCTGAGTTACGTGATCTAGTGAAGGAGCGCAGAGAAAGTATATTACAAAAGGAGACAGAA 1269

QY	1368	GGTTGAAAGGGGAAAGCAATACGGAAAGAGAGAGATTTTAGGAAATATTAGTCAGGCCCC	1427
		:::	
DB	363	alleThrGluValProProLys-----ArgGlnGlyLysGlyGlyPr	377
		: ::: :::	
QY	1428	TAGACAGTCAGGGAACCTGGGCAATAGGACCCACTGCACAAGGACCAATGTCATATTG	1487
		::: :::	
DB	377	oGlyLysGluIleGlnProGlyCysGlnSer---LeuGlnArgAsnGlnCysAlaIyrCy	396
		: ::: :::	
QY	1488	TAAAGAAAGAGGACACTGGGCAAGGAAGTGGCCCAAG---AAGGGAAACAAAGGACCAAG	1544
		: ::: :::	
DB	396	sLysGluIleGlyHisIleTrpLysAsnLysCysProGlnLeuLysGlyLysGlnGlyAsp--	415
		: ::: :::	
QY	1545	GATCCTAGTCTCTAGAAGAAGATTAAGAT	1572
		:::	
DB	416	-----SerGluGlnGluAspProAsp	422

RESULT 9

```

US-10-450-763-36834
; Sequence 36834, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36834
; LENGTH: 1577
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (472)..(488)
; OTHER INFORMATION: Eukaryotic and viral aspartyl proteases protein domain
; OTHER INFORMATION: identified by eMATRIX, accession number BL00141A, p-value=4.150e-
; OTHER INFORMATION: 10, raw score of 12.10
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (156)..(361)
; OTHER INFORMATION: Gag P30 core shell protein domain identified by PFam,
; OTHER INFORMATION: accession name Gag_p30, E-value=1.2e-67, PFam score of 238.2
US-10-450-763-36834

```

Alignment Scores:

Pred. No.	1.25e-32	Length:	1577
Score:	553.00	Matches:	151
Percent Similarity:	48.6%	Conservative:	64
Best Local Similarity:	34.2%	Mismatches:	149
Query Match:	19.4%	Indels:	78
DB:	5	Gaps:	19

US-10-723-552-3 COPY 585 2156 (1-1572) * US-10-450-763-36834 (1-1577)

298	QY	TGGCTGAATAAGCCAGAAAGACCGAGTCCCGGAATTC	TGCTCTTCTGGAGAGAAACAA	357
64	Db	TpLeuAsnThrGlyLysGlu-	-----LeuAlaLeuGlyValTrpThrPhe	78
358	QY	CACCTCGGCTGAAAAGTCAAGCCC-----TCTCCT	-----CATATCTACCCCGAG	402
		:::		
79	Db	GlutThrTrpGluArgGlySerProLeuLeuSerProArgGlyLeuHis	-----LeuHis	96
403	QY	ATTGAGGAGCACCGGCTTGCGCGGAACCCCAATCTGT	TCTCCCAACCCCTTATCTGGCA	462
		:::		
97	Db	-----LysThrAsnileProArgProArgValGly	-----	107

```
Db 426 AsnLysCysProGlnLeuLysGlyLysGlnGlyAsp-----SerGluGlnGluAsp 442
QY 1567 AAAGAT 1572
Db 443 ProAsp 444

RESULT 10
US-10-450-763-57482
; Sequence 57482, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 57482
; LENGTH: 1577
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (472)..(488)
; OTHER INFORMATION: Eukaryotic and viral aspartyl proteases proteins domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00141A, p-value=4.150e-
; OTHER INFORMATION: 10, raw score of 12.10
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (156)..(361)
; OTHER INFORMATION: (156)..(361)
; OTHER INFORMATION: Gag p30 core shell protein domain identified by PFam,
; OTHER INFORMATION: accession name Gag_p30, E-value=1.2e-67, Pfam score of 238.2
US-10-450-763-57482

Alignment Scores:
Pred. No.: 1-25e-32 Length: 1577
Score: 553.00 Matches: 151
Percent Similarity: 48.6% Conservative: 64
Best Local Similarity: 34.2% Mismatches: 149
Query Match: 19.4% Indels: 78
DB: 5 Gaps: 19

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-450-763-57482 (1-1577)
QY 298 TGGCTGAATAAGCCAAAGAACGAGTCCCGGAATTCGGCTCTTGGAGAGAAAAACAAA 357
Db 64 TrpLeuAsnThrGlyLysGlu-----LeuAlaLeuGlyValTrpThrPhe 78
QY 358 CACTCGGCTGAAAAGTCAAGCC-----TCTCCT-----CATATCTACCCCGAG 402
Db 79 GluThrTrpGluArgLysProLeuLeuSerProArgLysLeuHis-----96
QY 403 ATTGAGGAGCCACCGGCTTGGCGGGAACCCCAATCTGTTCCCAACCCCTTATCTGGCA 462
Db 97 -----LysThrAsnIleProArgProProArgValGly 107
QY 463 CAG-----GFTGCCGCGAGGGACCTTTGCCCCCTCTCTGGAGCTCCCGCGTGGAGGA 516
Db 108 LysArgGlyGlyGluAlaSerGlyGlu---ThrProPro-----119
QY 517 CCTGCTGACGAGCTCGAGCGCGAGGGCGGCCACCCCGAGCGGACAGACGAGATCGCG 576
Db 120 LeuAlaAlaGlyLeuArgProLysThrGly-----Ile 130
QY 577 ACATTACCGCTGCGCACG-----TACGGCCCTCCACACACCGGGGGGCCAATTG-----624
```

```
Db 131 GlnMetProLeuArgGluGlnTrpThrGlyIleAspGluAspGlyHisMetValGlu 150
QY 625 ---CAGCCCTCCAGTATTGGCCCTTTTCTCTCGAGATCTCTATAATTGAAACTAAC 681
Db 151 ArgArgValPheValTyrglnProPhePheSerAlaAspLeuLeuAsnTrpLysAsnAsn 170
QY 682 CATCCCTTTCTCGAGGATCCCAACGCTCACGGGTTGGTGGAGTCCCTTATGTTTC 741
Db 171 ThrProSerTyrglnGluLysProGlnAlaLeuIleAspLeuLeuGlnIleIleGln 190
QY 742 TCTCACCAGCTACTTGGGATTTGTCAACAGCTGCTGCAGACACTCTTTCACACCGAG 801
Db 191 ThrHisAsnProThrTrpAlaAspCysHisGlnLeuLeuMetPheLeuPheAsnThrAsp 210
QY 802 GAGCAGAGAGAAATCTATTAGAGCTAGAAAA-----AATGTTCTCGGGGCC 849
Db 211 GluArgArgValLeuGlnAlaAlaThrLysTrpLeuGluGluHisAlaProAlaAsp 230
QY 850 GACGGCGACCCACGCGGTTGCAAAATGAGATTGACATGGGATTTCCCTTAACTCGCC 909
Db 231 TyrGlnAsnPro-----GlnGluTyrglnValArgThrGlnLeuProGlyThrAspPro 247
QY 910 GGTGGGACTACAAACACGCTGAAGTAGGAGAGCTTTGAAAAATCTATCGCAGGCTCTG 969
Db 248 GlnTrpAspProAsnGluArgGluAspMetGlnArgLeuAsnArgTyrglnGluAlaLeu 267
QY 970 GTGGCGGCTCTCGGGGGCCCTCAAGACGGCGCCACTAATTTGGCTAAGTAAAGAGATG 1029
Db 268 LeuGluGlyLeuLysArgGlyAlaGlnLysAlaThrAsnValAla-----282
QY 1030 ATGACGAGACCGAATGAACCCCTCTGTTTTCTTGAGAGGCTCTTGGAGCCCTTCAGG 1089
Db 283 IleGlnGlyLysGluGluSerProAlaGlnPheTyrglnArgLeuGlnGlyAlaTyrgln 302
QY 1090 CGGTACACCCCTTTGATCCACCTTCAGAGCGCCCAAAAGCTCAGTGGCTTTGGCCCTTT 1149
Db 303 MetTyrglnProPheAspProAsnSerProGluAsnGlnArgMetIleAsnMetAlaLeu 322
QY 1150 ATAGACAGTCAAGCTTGGATATTAGAAAGCTTCAGACACTGGAAGGGTTACAGAG 1209
Db 323 ValSerGlnSerAlaGluAspIleArgArgLysLeuGlnLysGlnAlaGlyPheAlaGly 342
QY 1210 GCTGAGTTACGTGATCTAGTGAAGGAGGACAGAAAGTATATTACAAAGGGAGACAGAA 1269
Db 343 MetAsnThrSerGlnLeuLeuGluIleAlaAsnGlnValPheValAsnArgAspAlaVal 362
QY 1270 GAAGAAAGGGAAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1329
Db 363 SerProLysGluAsnArgArgGlu-----AsnGluHisGlnValArgArgAsn---378
QY 1330 CGGCAAGAGAAAGATTTGACTAAGATCTTGGCTGAGTGGTTGAAGGAAAGCAATACG 1389
Db 379 -----AlaAspLeuLeuAlaAlaIleThrGlyValProProLys 392
QY 1390 GAAAGAGAGAGAGATTTTAGGAAAATTAAGTCCAGCCCTAGACAGTCAAGGAAACCTGG 1449
Db 393 -----ArgGlnGlyLysGlyProGlyLysGluIleGlnProGly 406
QY 1450 AATAGGACCCCACTCGACAAGCAATGTGCATATTGTAAGAAAGAGGACACTGGGCA 1509
Db 407 CysGlnSer---LeuGlnArgAsnGlnCysAlaTyrglnGlyGluIleGlyHisTrpLys 425
QY 1510 AGGAACCTGCCCAAG---AAGGAAACAAAGCAAGGATCTAGTCTAGAGAGAT 1566
Db 426 AsnLysCysProGlnLeuLysGlyLysGlnGlyAsp-----SerGluGlnGluAsp 442
QY 1567 AAAGAT 1572
Db 443 ProAsp 444

RESULT 11
US-10-670-695-32
```


; Sequence 32, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Feline leukemia virus
US-10-670-695-32

Alignment Scores:
Pred. No.: 5,21e-32 Length: 611
Score: 543.50 Matches: 114
Percent Similarity: 67.0% Conservative: 30
Best Local Similarity: 53.0% Mismatches: 46
Query Match: 19.1% Indels: 25
DB: 4 Gaps: 4

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-670-695-32 (1-611)

QY 940 GAGAGCTTGAATCTATCGCAGGCTCTGTGGGGGTCTCCGGGGCGCTCAAGACGG 999
DB 2 GluHisLeuArgLeuTyrArgGlnLeuLeuLeuAlaGlyLeuArgGlyAlaAlaArgHis 21
QY 1000 CCCACTAATTTGGCTTAAGTAAGAGAGTGTAGCAGGACCGGAATGAACCCCTCTGTT 1059
DB 22 ProThrAsnLeuAlaGlnValLysGlnPheLeuGlnGlyLysGluThrProAlaSer 41
QY 1060 TTTCTTCAGAGGCTCTTGGAGCCTTACGGGGTACACCCCTTTGATCCCACTCAGAG 1119
DB 42 PheLeuGluArgLeuLysGluAlaTyrArgMetTyrThrProTyrAspProGluAspPro 61
QY 1120 GCCCAAAAGCCTCAGTGGCTTTGGCCCTTATAGCAGAGTGTAGCAGGCTTGGATATTAGAAG 1179
DB 62 GlyGlnAlaAlaSerValIleLeuSerPheIleTyrGlnSerSerProAspIleArgAsn 81
QY 1180 AAGCTTCAGAGACTGGAGGTTTACAGAGGCTGAGTGTACTGTAGTGAAGAGGCA 1239
DB 82 LysLeuGlnArgLeuGluGlyLeuGlnGlyPheThrLeuSerAspLeuLeuLysGluAla 101
QY 1240 GAGAAAGTATATTACAAAGGGACACAGAAAGAGGGAACAAGAAAGAGAGAGAA 1299
DB 102 GluLysIleTyrAsnLysArgGluThrProGluGluArgGluGluArg----- 117
QY 1300 AGACAGAAAGGGAGGAAGACGTAAATAAGCGCAAGAGAGAATTTGACTAAGATCTTG 1359
DB 118 LeuTrpGlnArgGlnGluArgAspLysLysArgHisLysGluMetThrLysValLeu 137
QY 1360 GCTGCAGTGGTGAAGGMAAAGCAATACGGAAGAGAGAGAGA-----GATTTT 1407
DB 138 AlaThrValValAlaGlnAsnArgAspLysAspArgGluGluSerLysLeuGlyAspGln 157
QY 1408 AGGAAATAGGTGAGCCCTTAGACAGTACAGGAACTGGGCAATAGACCCCACTGCAC 1467
DB 158 ArgLysIle-----ProLeuGly 163
QY 1468 AAGGACCAATGTGCATATTGTAAAGAGAGGACACTGGGCAAGGAACTGCCCAAGAG 1527
DB 164 LysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpValArgAspCysProAsnArg 183
QY 1528 GGAAACAAAGGACCA-----AGGATCCTAGCTCTAGAAGAA 1563

DB 184 ProArgLysLysProAlaAsnSerThrLeuLeuAsnLeuGluAsp 198
RESULT 12
US-10-450-763-36843
; Sequence 36843, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36843
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (391)..(407)
; OTHER INFORMATION: Eukaryotic and viral aspartyl proteases proteins domain
; OTHER INFORMATION: identified by eMATRIX, accession number BL00141A, p-value=4.150e-12
; OTHER INFORMATION: 10, raw score of 12.10
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (75)..(280)
; OTHER INFORMATION: gag p30 core shell protein domain identified by PFam,
; OTHER INFORMATION: accession name gag_p30, E-value=1.2e-67, PFam score of 238.2
US-10-450-763-36843
Alignment Scores:
Pred. No.: 6,72e-32 Length: 878
Score: 542.50 Matches: 139
Percent Similarity: 50.4% Conservative: 58
Best Local Similarity: 35.5% Mismatches: 137
Query Match: 19.0% Indels: 57
DB: 5 Gaps: 15
US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-450-763-36843 (1-878)
QY 436 TCTGTTCCCCACCCCTTATCTGGCACAG-----GGTGGCGGAGGGGACCCCTTTGCC 489
DB 18 AsnIleProArgProProArgValGlyLysArgGlyGlyGluAlaSerGlyGlu---Thr 36
QY 490 CTCCTCGAGCTCCGGCGGTGGAGGACCTGTCTCAGGAGCTCGAGCCGAGCGGGGCC 549
DB 37 ProPro-----LeuAlaAlaGlyLeuArgProLysThrGly--- 48
QY 550 ACCCGGAGCGGACAGACGAGATCGCGACATTACCGCTGGCAGC-----TACGGCCCT 603
DB 49 -----IleGlnMetProLeuArgGluGlnTrpTyrThrGly 60
QY 604 CCCACACCGGGGGCCCAATTG-----CAGCCCTCCAGTATTGGCCCTTTCTTCT 654
DB 61 IleAspGluAspGlyHisMetValGluArgValPheValTyrGlnProPhePheSer 80
QY 655 GCAGATCTCTATATTGCGAAACTTAACCATCCCTTTCTCGGAGGATCCCAACGCGCTC 714
DB 81 AlaAspLeuLeuAsnTrpLysAsnAsnThrProSerTyrThrGluLysProGlnAlaLeu 100
QY 715 ACGGGGTGTGGAGTCCCTTATGTTCTCTCACCAGCTACTTGGGATGATTGTCAACAG 774
DB 101 IleAspLeuLeuGlnThrIleIleGlnThrHisAsnProThrTrpAlaAspCysHisGln 120
QY 775 CTGCTGCAGACACTCTTTCACAAACCGAGGAGCGAGAGAAATTCATTATTAGAGCTAGAAA 834

Db	109	---	LeuProProSerAlaProSerLeuProLeuGluProProLeuSerThrProProArg	121
Qy	418	-----	-----GCTTGGCCGGAACCCCAA-----	435
Db	128	SerSerLeuTyrProAlaLeuThrProSerLeuGlyAlaLeuProLeuGluProGlnValLeu	147	
Qy	436	-----	-----TCTGTTCCCCACCCCTTTATCTGTGCA	462
Db	148	SerAspSerGlyProLeuIleAspLeuLeuThrGluAspProProTyr	165	
Qy	463	CAGGTCGCGGAGGGGACCCCTTCCCTCCCTCT	495	
Db	166	-----	-----ArgAspProArgProProSerAspArgAspGlyAsnGlyGlyGlu	181
Qy	496	GGAGCTCCGCGCTGGAGGACCT-----	540	
Db	182	AlaThrProAlaGlyGluAlaProAspProSerProMetAlaSerArgLeuArgGlyArg	201	
Qy	541	AGGGCGCCACCCCGGAGCGACAGACGAGATCGCGACATTACCGCTCGCGACGTACGCG	600	
Db	202	ArgGluProProValAlaAspSerThrThrSerGlnAlaPheProLeuArgThrGlyGly	221	
Qy	601	CCTCCACACCGGGGGGCAATTGCAGCGCCCTCCAGATTGGCCCTTTTCTTCTGCAGAT	660	
Db	222	-----	-----AsnGlyGln-----LeuGlnTyrTrpProPheSerSerSerAsp	234
Qy	661	CTCTATAATTGGAAACCTAAC-----	696	
Db	235	LeuTyrIleThrProValAsnSerLeuGluIysHisSerTrpTyrHisGlyProValSer	254	
Qy	697	GAGGAT 702		
Db	255	ArgAsn 256		
RESULT 15				
US-10-732-923-13455				
; Sequence 13455, Application US/10732923				
; Publication No. US20050108791A1				
; GENERAL INFORMATION:				
; APPLICANT: Edgerton, Michael D				
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES				
; FILE REFERENCE: 38-15(52796)C				
; CURRENT APPLICATION NUMBER: US/10/732,923				
; CURRENT FILING DATE: 2003-12-10				
; PRIOR APPLICATION NUMBER: 10/310,154				
; PRIOR FILING DATE: 2002-12-04				
; NUMBER OF SEQ ID NOS: 24149				
; SEQ ID NO 13455				
; LENGTH: 981				
; TYPE: PRT				
; ORGANISM: Abelson murine leukemia virus				
US-10-732-923-13455				
Alignment Scores:				
Pred. No.:		1.8e-28	Length:	981
Score:		497.00	Matches:	116
Percent Similarity:		48.9%	Conservative:	22
Best Local Similarity:		41.1%	Mismatches:	70
Query Match:		17.5%	Indels:	74
DB:		5	Gaps:	10
US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-732-923-13455 (1-981)				
Qy	1	ATGGGACAGAGGTGACGACCCCTTAGTTAGTCTCGACCATTTGGACTGAAGTTAAA	60	
Db	1	MetGlyGlnThrValThrThrProLeuSerLeuThrLeuGlyHisTrpIysAspValGlu	20	
Qy	61	TCCAGGCGCTCATATTTGTTCAGTTTCAGGTTAAGAGGACCTTGGCAGACTTCTGTGTC	120	
Db	21	ArgIleAlaHisAsnGlnSerValAspValIysIysArgArgTrpValThrPheCysSer	40	
Qy	121	TCTGAATGGCGACATTCGATGTGTGGATGGCCATCAGAGGGAGCCTTTAATTCAGATT	180	

```
Db 41 AlaGluTrpProThrPheAsnValGlyTrpProArgAspGlyThrPheAsnArgAspLeu 60
Qy 181 ATCTGGCTGTTAAAGCAGTTATTTTTCAGACTGGACCGGCTCTCATCCCGATCAGGAG 240
Db 61 IleThrGlnValLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
Qy 241 CCCTATATCTTACGTGGCAGATTGGCAGAGGATCTCCGCCCATGGTTAAACCATGG 300
Db 81 ProTyrIleValThrTrpGluAlaLeuAlaPheAspProProTrpValIysProPhe 100
Qy 301 CTGAATAAGCCAAAGACCGAGTCCCGCAATTCTGGCTCTTGGAGAGAAAAACAACAC 360
Db 101 ValHis---ProLysProProPro----- 108
Qy 361 TCGGCTGAAAAGTCAAGCCCTCTCCTCATATCTACCCCGAGATTGAGGAGCCACCG--- 417
Db 109 ---LeuProProSerAlaProSerLeuProLeuGluProProLeuSerThrProProArg 127
Qy 418 -----GCTTGGCCGGAACCCCAA----- 435
Db 128 SerSerLeuTyrProAlaLeuThrProSerLeuGlyAlaLysProLysProGlnValLeu 147
Qy 436 -----TCTGTTCCCCACCCCTTATCTGGCA 462
Db 148 SerAspSerGlyGlyProLeuIleAspLeuLeuThrGluAspProProTyr----- 165
Qy 463 CAGGTGCCGCGAGGGGACCTTTGCCCTCCT----- 495
Db 166 -----ArgAspProArgProProSerAspArgAspGlyAsnGlyGlyGlu 181
Qy 496 GGAGCTCCGGCGGTGGAGGACCT-----GCTGCAGGACTCGGAGCCGG 540
Db 182 AlaThrProAlaGlyGluAlaProAspProSerProMetAlaSerArgLeuArgGlyArg 201
Qy 541 AGGGCGCCACCCCGAGCGGACACGAGATCGCGACATTACCGCTCGGCACGTACCGC 600
Db 202 ArgGluProProValAlaAspSerThrThrSerGlnAlaPheProLeuArgThrGlyGly 221
Qy 601 CCTCCACACCGGGGGCCAAATTGAGCCCTCCAGTATTGGCCCTTTTCTTCTGCAGAT 660
Db 222 -----AsnGlyGln-----LeuGlnTyrTrpProPheSerSerSerAsp 234
Qy 661 CTCTATAATTGAAAACCTAAC-----CATCCCCCTTTCTCG 696
Db 235 LeuTyrIleThrProValAsnSerLeuGluLysHisSerTrpTyrHisGlyProValSer 254
Qy 697 GAGGAT 702
Db 255 ArgAsn 256
```

Search completed: February 14, 2006, 17:17:27
Job time : 223.636 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:19:01 ; Search time 1.37849 Seconds
(without alignments)
2992.933 Million cell updates/sec

Title: US-10-723-552-3_COPY_585_2156
Perfect score: 2848
Sequence: 1 ATGGGACAGCGTGACGAC.....CTCTAGAGAGATAAAGAT 1572

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 194028

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10723552/runat 14022006 125151 13368/app query.fasta_1
-DB=Published_Applications AA New -QFMT=fastan -SURFIX=rapbn -MINMATCH=0.1
-LOOPCUT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-MAXLEN=2000000000 -HOST=abs02p
-USR=US10723552 @CGN 1 1 17 @runat 14022006 125151 13368 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DRV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRAUS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pap:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1358	47.7	537	7	US-11-129-442-47
2	163.5	5.7	416	7	US-11-226-657-136
3	138.5	4.9	1388	6	US-10-821-234-1143
4	136.5	4.8	538	7	US-11-124-368A-311
5	136.5	4.8	538	7	US-11-124-368A-312
6	136.5	4.8	538	7	US-11-124-368A-313
7	136.5	4.8	1560	7	US-11-059-982-1
8	132	4.6	1410	6	US-10-821-234-1050
9	129.5	4.5	2760	7	US-11-124-367A-444

10	129.5	4.5	2803	7	US-11-124-367A-442	Sequence 442, App
11	129.5	4.5	2803	7	US-11-124-367A-445	Sequence 445, App
12	129.5	4.5	2984	7	US-11-124-367A-443	Sequence 443, App
13	129.5	4.5	3027	7	US-11-124-367A-441	Sequence 441, App
14	129	4.5	1744	7	US-11-182-016-22	Sequence 22, Appl
15	128.5	4.5	615	6	US-10-982-545-14	Sequence 14, Appl
16	128	4.5	948	6	US-10-523-477-14	Sequence 14, Appl
17	127.5	4.5	716	7	US-11-150-845-16	Sequence 16, Appl
18	127.5	4.5	716	7	US-11-150-847-16	Sequence 16, Appl
19	127	4.5	638	7	US-11-150-845-24	Sequence 24, Appl
20	127	4.5	638	7	US-11-150-847-24	Sequence 24, Appl
21	127	4.5	689	7	US-11-024-959-510	Sequence 510, App
22	125.5	4.4	616	6	US-10-982-545-5	Sequence 5, Appl
23	123	4.3	639	6	US-10-821-234-907	Sequence 907, App
24	123	4.3	863	7	US-11-169-041-167	Sequence 167, App
25	123	4.3	915	6	US-10-821-234-1514	Sequence 1514, Ap
26	121.5	4.3	8746	7	US-11-098-686-10232	Sequence 10232, A
27	121	4.2	537	6	US-10-504-364-6	Sequence 6, Appl
28	121	4.2	537	6	US-10-504-364-7	Sequence 7, Appl
29	120.5	4.2	551	6	US-10-821-234-1580	Sequence 1580, Ap
30	120.5	4.2	551	6	US-10-504-364-1	Sequence 1, Appl
31	120.5	4.2	551	6	US-10-504-364-2	Sequence 2, Appl
32	120.5	4.2	551	6	US-10-504-364-5	Sequence 5, Appl
33	119.5	4.2	715	7	US-11-150-845-14	Sequence 14, Appl
34	119.5	4.2	715	7	US-11-150-847-14	Sequence 14, Appl
35	119.5	4.2	1823	6	US-10-220-824-8	Sequence 8, Appl
36	119.5	4.2	1823	6	US-10-995-561-988	Sequence 988, App
37	119.5	4.2	2157	6	US-10-995-561-991	Sequence 991, App
38	119	4.2	609	7	US-11-150-845-20	Sequence 20, Appl
39	119	4.2	609	7	US-11-150-847-20	Sequence 20, Appl
40	119	4.2	891	7	US-11-182-016-38	Sequence 38, Appl
41	119	4.2	2102	6	US-10-995-561-990	Sequence 990, App
42	119	4.2	2108	6	US-10-995-561-989	Sequence 989, App
43	117.5	4.1	915	6	US-10-995-561-1003	Sequence 1003, Ap
44	117.5	4.1	917	6	US-10-995-561-1000	Sequence 1000, Ap
45	117.5	4.1	940	6	US-10-995-561-1004	Sequence 1004, Ap

ALIGNMENTS

RESULT 1
US-11-129-442-47
; Sequence 47, Application US/11129442
; Publication No. US2006002951A1
; GENERAL INFORMATION:
; APPLICANT: Kleiman, Lawrence
; APPLICANT: CEN, Shan
; APPLICANT: GUO, Fei
; TITLE OF INVENTION: Inhibition of the tRNAlys3-Primed Initiation of Reverse
; TITLE OF INVENTION: Transcription in HIV-1 by APOBEC3G
; FILE REFERENCE: 11169.257
; CURRENT APPLICATION NUMBER: US/11/129,442
; PRIOR FILING DATE: 2005-05-16
; PRIOR APPLICATION NUMBER: CA 2,467,312
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Murine leukemia virus
US-11-129-442-47

Alignment Scores:
Pred. No.: 3.92e-88 Length: 537
Score: 1358.00 Matches: 277
Percent Similarity: 62.7% Conservat: 76
Best Local Similarity: 49.2% Mismatches: 142
Query Match: 47.7% Indels: 68
DB: 7 Gaps: 12

US-10-723-552-3_COPY_585_2156 (1-1572) x US-11-129-442-47 (1-537)

```
QY 1 ATGGACAGACGGTCACGACCCCTCTAGTTGACTCTCGACCATTTGACGTCAAGTTAAA 60
Db |||||||||||||||||||:|:|:| ||||||||||||||| ||||||| |||||||
1 MetGlyGlnThrValThrProLeuSerLeuThrLeuGluHisTrpGlyAspValGln 20
QY 61 TCCAGGGCTCATAAATTTGTCAGTTTCAAGTAAAGAGGACCTTGGCAGACTTCTGTGTC 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 ArgIleAlaSerAsnGlnSerValAspValLysLysArgTrpValThrPheCysSer 40
QY 121 TCTGAATGGCCGACATTCATGTTGGATGGCCATCAGAGGAGACCTTAAATCTGAGATT 180
Db |||||||:|:|:| |||||:|:|:| |||||:|:|:| |||||:|:|:| |||||:|:|:|
41 AlaGluTrpProThrPheGlyValGlyTrpProGlnAspGlyThrPheAsnLeuAspIle 60
QY 181 ATCTGGCTGTTAAAGCAGTATTTTTCAGACTGGACCGGCTCTCATCCCGATCAGGAG 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 IleLeuGlnValLysSerLysValPheSerProGlyProHisGlyHisProAspGlnVal 80
QY 241 CCTATATCTTACCTGTCGACAGATTGGCAGAGGATCTCGCCCATGGTTAAACCATGG 300
Db |||||||:|:|:| |||||:|:|:| |||||:|:|:| |||||:|:|:| |||||:|:|:|
81 ProTyrIleValThrTrpGluAlaIleAlaTyrGluProProTrpValLysProPhe 100
QY 301 CTGAATAAGCCAGAAAGCCAGGTCCTCGA-----ATTCTGGCTCTTGGAGAGAAAAC 354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 ValSerProLysLeuSerProSerProThrGlyProIleLeuProSerGlyProSerThr 120
QY 355 AAACACTCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCCGAGATTGAGGAGCCA 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 GlnProProProArgSerAlaLeu---TyrProAlaLeuThrProSerIleLysProArg 139
QY 415 CCGGCTTGGCCGGAACCCCAATCT-----GTT 441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
140 ProSerLysProGlnValLeuSerAspAsnGlyGlyProLeuIleAspLeuLeuThrGlu 159
QY 442 CCCCACCCCTTATCTGGCACAGGGTCCCGAGGAGACCTTTCGCCCTCTCGA--- 498
Db ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
160 AspProProProTyr-----GlyGluGlnGlyProSerSerSerAspGlyAsp 175
QY 499 -----GCTCCGGCGGTGGAGGACCTCTGCA 525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 GlyAspArgGluGluAlaThrSerThrSerGluIleProAlaProSerProMetValSer 195
QY 526 GGGACTCGAGCGCGGAGGGGCCACCCCGGAGCGGACAGACAGATCGGACATTACCG 585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 ArgLeuArgGlyLysArgAspProProAlaAlaAspSerThrThrSerArgValPhePro 215
QY 586 CTGCGCAGCTAGCGGCTCCACACCGGGGGGCCAAATTGCGAGCCCTCCAGTATTTGGCCC 645
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 LeuArgLeuGlyGly-----AsnGlyGln-----LeuGlnTyrTrpPro 228
QY 646 TTTTCTCTCGAGATCTCTATAATTGGAATACTAACCATCCCTTTCTCGAGGATCCC 705
Db ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
229 PheSerSerSerAspLeuTyrAsnTrpLysAsnAsnProSerPheSerGluAspPro 248
QY 706 CAACGCCTCAGCGGCTTGTGAGTCCCTTATGTTCTCTCACAGCCTACTTGGGATGAT 765
Db ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 GlyLysLeuThrAlaLeuIleGluSerValLeuThrHisGlnProThrTrpAspAsp 268
QY 766 TGTCAACAGCTGCTCAGACACTTTCACACCGGAGGCGAGAGAGAAATCTATAGAG 825
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 CysGlnGlnLeuLeuGlyThrLeuLeuThrGlyGluGluLysGlnArgValLeuLeuGlu 288
QY 826 GCTGAAAAAATGTTCTTGGGCGCCGCGGCGGACCGGCTGAAAGGTAGGAGAGC 885
Db ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
289 AlaArgLysAlaValArgGlyAsnAspGlyArgProThrGlnLeuProAsnGluValAsn 308
QY 886 ATGGATTTCCTTAACCTCGCCCGGTTGGGACTACAAACAGCGCTGAAAGGTAGGAGAGC 945
Db ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
309 SerAlaPheProLeuGluArgProAspTrpAspTyrThrProGluGlyArgAsnHis 328
QY 946 TTGAAAAATCTATGCGCAGCTCTGTGGCGGCTCTCGGGGCGCTCAAGACGGCCCACT 1005
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
329 LeuValLeuTyrArgGlnLeuLeuAlaGlyLeuGlnAsnAlaGlyArgSerProThr 348
QY 1006 AATTGGCTTAGGTAAGAGACTGATGAGGGAGCCGAATGAACCCCTCTCTTTTCTT 1065
```

```
Db |||||||||||||||||||:|:|:| ||||||||||||||| ||||||| |||||||
349 AsnLeuAlaLysValLysGlyIleThrGlnGlyProLeuGlnSerProSerAlaPheLeu 368
QY 1066 GAGAGGCTCTTTGGAGCCTTCAGCGGTACACCCCTTTTGATCCACCTCAGAGGCCCAA 1125
Db ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 GluArgLeuLysGluAlaTyrArgArgTyrThrProTyrAspProGluAspProGlyGln 388
QY 1126 AAAGCCTCAGTCGGCTTGGCCCTTTATAGGACAGTCAGCCTTGGATATTAGAAAGAGCTT 1185
Db ||:|:|:| |||||:|:|:| |||||:|:|:| |||||:|:|:| |||||:|:|:|
389 GluThrAsnValSerMetSerPheIleTrpGlnSerAlaProAspIleGlyArgLysLeu 408
QY 1186 CAGAGACTGGAGGTTTACAGAGGCTGAGTTACGTGATCTAGTCAAGGAGCGCAGAGAAA 1245
Db ||:|:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
409 GluArgLeuGluAspLeuLysSerLysThrLeuGlyAspLeuValArgGluAlaGluArg 428
QY 1246 GTATATTACAAAAGGAGACAGAGAAAGGAAAGGAAAGAAAGAGAGAGAGAGAGAG 1305
Db ||:|:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 IlePheAsnLysGlyGluThrProGluGluArgGluGluArgValArgGluThrGlu 448
QY 1306 CAAAGGAGGAGAAAGA-----CGTAATAAACGG 1332
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
449 GluLysGluGluArgArgArgAlaGluGluGlnLysGluLysGluLysArgAspArg 468
QY 1333 CAAGAGAGAATTTGACTTAAGATCTTGCTGCAGTGGTTGAAGGAGAAAGCAATACGAA 1392
Db ||:|:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
469 ArgHisArgGluMetSerLysLeuAlaThrValValSerGlyGln----- 484
QY 1393 AGACAGAGAGATTTTAGGAAAATTTAGTTCAGCCCTTAGACAGTCAGGGAACCTGGGCAAT 1452
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
485 -----ArgGlnAspArgGlnGlyGluArg-----Arg 494
QY 1453 AGGACCCACTCGACAGGACCAATGTGCATTTGTAAGAAAGAGGACACTGGCAGG 1512
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 ArgProGlnLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaLys 514
QY 1513 AACTGCCCAAGAAG-----GGAAACAAAGGACCAAG-----ATCCTAGCT 1554
Db ||:|:|:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
515 AspCysProLysLysProArgGlyProArgGlyProArgProGlnThrSerLeuLeuThr 534
QY 1555 CTAGAAGAA 1563
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
535 LeuAspAsp 537

RESULT 2
US-11-226-657-136
; Sequence 136, Application US/11226657
; Publication No. US20060025574A1
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006G13APIC1D2
; CURRENT APPLICATION NUMBER: US/11/226,657
; CURRENT FILING DATE: 2005-09-15
; PRIOR APPLICATION NUMBER: 10/062,831
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/690,454
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US98/10868
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: 1997-05-30
```



```

Db      229 SerAspSerHisProArgLysLysLysLysLysGluLysLysGluAspLePheArgGpHe 248
Qy      1096 -----ACCCCTTTTGATCCCACTCAGAGGCCCAAAGAAGCCTCAGTGGCTTTGGCGCTTT 1149
Db      249 ProValAlaProLeulleProTyPrProLeulleThrLysGluAspIleAenAlalleGlu 268
Qy      1150 ATAGGACAGTCAGCCTTGGAATATT-----AGA 1176
Db      269 MetGluGluAspLysArgAspLeulleSerArgGluIleSerLysPheArgAspThrHis 288
Qy      1177 AAGAGCTTCAGAGACTGGAGGGTTACAGGAGGCTGAGTTACGTGATCTAGTGAAG--- 1233
Db      289 LysLysLeuGluGluLysGlyLysLysGluLysGluLysGluLysGluLysGlu 308
Qy      1234 -----GAGGCAGAGAAAGTAGTATTACAAAAAGGAGAGACAGAGAAGAA 1275
Db      309 ArgArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 328
Qy      1276 AGGGAACAAGAAAAGAGAGAGAAAGAGAG--GAAGGGAGGAAAGACGTATAATAACGG 1332
Db      329 ArgGluArgGluArgGluArgGluLysGluLysGluLysGluLysGluLysGluArgGlu 348
Qy      1333 CAAGAGAGAATTTGACTAAGATCTTGGCTGCAGTGGTTCAAGGGGAAAGCAATACGGAA 1392
Db      349 ArgAspArgAspArgThrLys-----GluArgAspArgAspArgGlu 364
Qy      1393 AGAGAGAGAGATTTAGGAAAAATTAGTCAAGGCCCTAGACAGTCAGGAAACCTGGGCAAT 1452
Db      365 ArgAspArgAspArgAspArgGluArgSerSerAspArgAsnLysAspArgIleArgSer 384
Qy      1453 AGG 1455
Db      385 Arg 385

RESULT 3
US-10-821-234-1143
; Sequence 1143, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1143
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1143

Alignment Scores:
Pred. No.:          0.00943           Length:       1388
Score:             138.50            Matches:     119
Percent Similarity: 34.3%            Conservative: 69
Best Local Similarity: 21.7%         Mismatches:  159
Query Match:       4.9%              Indels:     201
DB:                6                 Gaps:        26

US-10-723-552-3_COPY_585_2156 (1-15172) x US-10-821-234-1143 (1-1388)
Qy      80 CAGTTCAGGTTAAGAGGACCTTGGCAGACTTCTGCTCTCTGAATGGCGACATTCG 139
Db      193 GlntyrThrArgLyslaGluePheArgLysLeuCysAspAsnLeuArgMetHisLeuSer 212
Qy      140 ATGTTGGATGGCCATCAGAGGGGACCTTTTAATTCCTGAGATTATTCCTGGCTGTTAAAGCAG 199

```

213 GlnIleGlnArgHisAsnGlnSerThrAlaIleAsnLeuAsnAsnProGluSerGln 232
200 TTATTTTCAGACTGGACCGGCTCTCATCCGAGTCAGAGCCCTATATCTTACGTGGC 259
233 ---SerMetHisLeuGluThrArgLeuValGlnLeuAspSerAlaIleSerMet----- 249
260 AAGATTTCGCGAGGATCTCCGCG-----CATGGTTAAACCATGCGTGA 304
250 GluLeuTrpGlnGluAlaPheLysAlaValGluAspIleHisGlyLeu----- 265
305 ATAGCCCAAGAAAGCCAGGTCCTCCGAATTCCTGGCTCTTGGAG-----AGAAAAACA 355
266 PheSerLeuSerLysProLysProLysProGlnLeuMetAlaAsnTyTyAsnLysVal 285
356 AACACTCGGCTGAAAAAGTCAGCCCTCTCTCATATCTACCCGAGATTGAGAGCCAC 415
286 SerThrValPheTrpLysSerGlyAsnAlaLeuPheHisAlaSerThrLeu-----His 303
416 CGGCTTGGCGGACCCCAATCTGTTCCTCCACCCCTTATCTGCGCAGGGTCCGCGA 475
304 ArgLeu-----TyrHisLeuSerArg 310
476 GGGACCCCTTGGCCCTCTCGAGCTCCGGGG-----TGGAGGACCTCTGCGAGGA 529
311 -----GluMetArgLysAsnLeuThrGlnAspGluMetGlnArg 323
530 CTCGAGCGCGGAGG-----GGCGCCACCCCGGAGGCGGACAG 565
324 Met--SerThrArgValLeuLeuAlaThrLeuSerIleProIleThrProGluArgThrA 343
566 AC-----GAGATCCGCGA 577
343 spIleAlaArgLeuLeuAspMetAspGlyIleIleValGluLysGlnArgArgLeuAla 363
578 CATTACCCCTCGCGACGTACGCGCCCTCCACACCGGGGGG-----C 619
363 hrLeu--LeuGlyLeuGlnAlaProProThrArgIleGlyLeuIleAsnAspMetValA 382
620 AATTCAGCCCTCCAGTATTGGCCCTTTCTTCTGCGAGATCTCTATATTTGG-----AAAA 676
382 rgPheAsnValLeuGlnIleValProGluValLysAspLeuTyAsnTrpLeuGluV 402
677 CTACCACTCCCTCTTCTCG-----G 697
402 alGluPheAsnProLeuLysLeuCysGluArgValThrLysValLeuAsnTrpValArg 422
698 AGGATCCC----- 705
422 luGlnProGluLysGluProGluLeuGlnGlnIleValProGlnLeuGlnAsnAsnThrI 442
706 -----CAACGCTCAGCGGTGGTGGAGTCCCTTATGTCTCTACACGAC 751
442 leLeuArgLeuGlnValSerGlnIleTyGlnSerIleGluPheSerArgLeu 462
752 CTACT-----TGGATGATTCACACGCTCTGCGACACACTCTTCACACCGAGG 802
462 hrSerLeuValProPheValAspAlaPheGlnLeu----- 473
803 AGCGAGAGAGATTCATTAGAGGCTAGAAAAATGTTCTCGGGCGCGAGCGGCGACCCA 862
474 -----GluArgAlaIleValAspAla-----AlaArgHisC 484
863 CGCGGTTGCAAAATGAGATTGACATGGGATTCCTTAACCTGCGCCCGGTGGGACTACA 922
484 ysAspLeuGlnValArgIleAspHisThrSerArgThrLeuSerPheGlySerAspLeuA 504
923 ACACGGCTCAGGTAGGAGGAGCTTGAAAAATCTATCGCCAGGCTCTGTGGCGGCTCTCC 982
504 snTyAlaThr--ArgGluAspAlaProIle----- 513
983 GGGGCGCTCAAGACGGGCCACTAATTTGGCTTAAGGTAAGAGAGTGTGCGAGGACCGA 1042
514 -----GlyProH 516

1043 ATGAACCCCTCTGTGTTTTCTTGAGAGCTCTTGGAGCGCTTACAGCGGTACACCCCTT 1102
516 is-----LeuGlnSerMet----- 520
1103 TTGATCCACCTCAGAGGCCCAAAAGCCTCAGTGGCTTTGGCCTTTATAGGACAGTCAG 1162
521 -----ProSerGluGlnIleArgAsnGlnLeuThrAlaMetSerValLeuAlaLysA 539
1163 CTTGGATATTAGAAAGCTTCCAGACTCTGGAAGGGTTACAGAG----- 1209
539 laLeuGluValIleLysProAlaHisIleLeuGlnGluLysGluGlnHisGlnLeuA 559
1210 -----GCTGACTACGTCTAGTGAAGGAGGACAGAAAGTATTACAAAAGGG 1261
559 laValThrAlaTyLeuLysAsnSerArgLysGluHisGlnArgIleLeuAlaArgA 579
1262 AGACAGAAGAAAGGAAACAAAGAAAGAG-----AGAGAAAGAGAGG 1306
579 lnThrIleGluLysGluArgLeuSerLeuAsnIleGlnArgGluLysGluG 599
1307 AAAGGAGGAAGA-----C 1321
599 luLeuGluGlnArgGluAlaGluLeuGlnLysValArgLysAlaGluGluArgLeuA 619
1322 GTAATAACGCAAGAGAGAAATTTGACTAAGATCTTGGCTGCAGTGGTTGAAGGAAAA 1381
619 rgGlnGluAlaLysGluArgLysGluArgIleLeuGlnGluHisGlnIleLysL 639
1382 GCAATACGAAAGAGAGAGA 1401
639 ysLysThrValArgGluArg 645

RESULT 4

US-11-124-368A-311
; Sequence 311, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 2112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 311
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-311

Alignment Scores:
Pred. No.: 0.0123 Length: 538
Score: 136.50 Matches: 83
Percent Similarity: 31.7% Conservative: 23
Best Local Similarity: 24.9% Mismatches: 112
Query Match: 4.8% Indels: 116
DB: 7 Gaps: 12

US-10-723-552-3_COPY_585_2156 (1-1572) x US-11-124-368A-311 (1-538)

301 CTGATAAGCCCAAGAACCCAGTCCCGCAATTCCTGGCTCTTGG----- 345
41 LeuGlyGlyProArgLysSerGlyProArgGlyAlaThrGlyGlyProGlyAspGluPro 60
346 -----GAGAAAAACAAACACTCGGCTGAGAAAGTCAAGCCCTCTCTCAT 390

```

Db      61 LeuGluProAlaArgGluGlnGlySerLeuAspAlaGluArgAsnGlnArgGlySerPhe 80
Qy      391 ATCTACCCCGAGATTGAG-----GAGCCACCG-----GCTTGGCGG 426
Db      81 GluAlaProArgTyrGluGlySerPheProAlaGlyProProThrArgAlaLeuPro 100
Qy      427 GAACCCCAATCTGTTCCCGCA-----CCCTTATCTGGGCACAGGGTCCCGCG 474
Db      101 LeuProGlnSerLeuProProAspPheArgLeuGluProThrAlaProAlaLeuSerPro 120
Qy      475 AGGGACCCCTTTGGC-----CCTCTGGAGCTCCGCGCGTG 510
Db      121 ArgSerSerPheAlaSerSerAlaSerAspAlaSerLysProSerPro-ArgG1 140
Qy      511 GAGGACCTGCTGCAGGACTCGGAGCGGAGGGCGCCACCCGCGAGCGACAGACGAG 570
Db      140 ySerLeuLeuLeuAspGlyAlaGlyAlaGlyAla-----GlyGlySerAr 156
Qy      571 ATCCGGACATTACCGCTGCGCAGCTAGCGCCCTC-----CCACACCGGGGGCG 618
Db      121 ArgSerSerPheAlaSerSerAlaSerAspAlaSerLysProSerPro-ArgG1 140
Qy      511 GAGGACCTGCTGCAGGACTCGGAGCGGAGGGCGCCACCCGCGAGCGACAGACGAG 570
Db      140 ySerLeuLeuLeuAspGlyAlaGlyAlaGlyAla-----GlyGlySerAr 156
Qy      571 ATCCGGACATTACCGCTGCGCAGCTAGCGCCCTC-----CCACACCGGGGGCG 618
Db      156 gProCysSerAsnArgThrSerGlyIleSerMetGlyTyrAspGlnArgHis-glySerP 176
Qy      619 CAATTGACGCCCTCCAGTATTGGCCCTTTCTTCGAGATCTCTATAATGGAAACT 678
Db      176 roLeuProAlaGlyProCysLeuPhe-----GAGCCACCG-----TCACCGACCTACTTGGAT 762
Qy      679 AACCATCCCTTCTTCGAGGATCCCAACGCCCTCACGGGGTGGTGGAGTCCCTTATG 738
Db      185 -----GlyProProLeuAlaGlyAlaProAlaGlyTyrSerProGlyGlyValProSerA 203
Qy      739 TTCTC----- 743
Db      203 laTyrProGluLeuHisAlaAlaLeuAspArgLeuTyrAlaGlnArgProAlaGlyPheG 223
Qy      744 -----TCACCGACCTACTTGGAT 762
Db      223 lyCysGlnGluSerArgHisSerTyrProProAlaLeuGlySerProGlyAlaLeuAlaG 243
Qy      763 GATTGTCAACAGCTGTGCAGACTCTTCAACACCGAGGAGCGAGA-----GAGAATT 816
Db      243 lyAlaGlyValGlyAlaAlaGlyProLeuGluArgArgGlyAlaGlnProGlyArgHis 263
Qy      817 CTATTAGAGCTAGAAAATGTTCTCTGGGCGGCGGAGCGGACCCACCGGTTTCAAAAT 876
Db      263 erValThrGlyTyrGlyAspCysAlaValGlyAlaArgTyrGlnAspGluLeuThr---- 281
Qy      877 GAGATTGCATGGATTCCCTTAACCTGCGCCCGGTTGGGACTACACACGCGCTGAAGT 936
Db      281 ----- 281
Qy      937 AGGAGAGCTTGAATAATCTATCGCAGGCTCTGGTGGCGGTCTCCGGGGCGCTCAAGA 996
Db      282 -----AlaLeuLeuArgLeuThrValGlyThrGlyGlyArgGlu-AlaGlyAlaArgGly 299
Qy      997 CGGCCCACTAATTGGCTAAGTAAAGAGAAGTGTGACGAGGACCGAATGAACCCCTCT 1056
Db      300 GluProSerGlyIleGluPro-----SerGlyLeuGluGluProProGly 314
Qy      1057 GTTTTCTTGAGGCTCTTGGAGCCCTTCAGGCGG 1092
Db      315 ProPheValProGluAlaAlaArgAlaArgMetArg 326

```

RESULT 5

```

US-11-124-368A-312
; Sequence 312, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof

```

```

; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-312

```

Alignment Scores:

```

Pred No.: 0.0123 Length: 538
Score: 136.50 Matches: 83
Percent Similarity: 31.7% Conservative: 23
Best Local Similarity: 24.9% Mismatches: 112
Query Match: 4.8% Indels: 116
DB: 7 Gaps: 12

```

```

US-10-723-552-3_copy_585_2156 (1-1572) x US-11-124-368A-312 (1-538)
Qy      301 CTGAATAAGCCCAAGAAAGCCAGGTCCCGCAATTCTGGCTCTTGA----- 345
Db      41 LeuGlyGlyProAlaArgGlySerGlyProArgGlyAlaThrGlyGlyProGlyAspGluPro 60
Qy      346 -----GAGAAAACAAACACTCGGCTGAAAAAGTCAAGCCCTCTCTCAT 390
Db      61 LeuGluProAlaArgGluGlnGlySerLeuAspAlaGluArgAsnGlnArgGlySerPhe 80
Qy      391 ATCTACCCCGAGATTGAG-----GAGCCACCG-----GCTTGGCGG 426
Db      81 GluAlaProArgTyrGluGlySerPheProAlaGlyProProThrArgAlaLeuPro 100
Qy      427 GAACCCCAATCTGTTCCCGCA-----CCCTTATCTTCGCGACAGGGTCCCGCG 474
Db      101 LeuProGlnSerLeuProProAspPheArgLeuGluProThrAlaProAlaLeuSerPro 120
Qy      475 AGGGACCCCTTTGGC-----CCTCTGGAGCTCCGCGCGTG 510
Db      121 ArgSerSerPheAlaSerSerAlaSerAspAlaSerLysProSerPro-ArgG1 140
Qy      511 GAGGACCTGCTGCAGGACTCGGAGCGGAGGGCGCCACCCGCGAGCGACAGACGAG 570
Db      140 ySerLeuLeuLeuAspGlyAlaGlyAlaGlyAla-----GlyGlySerAr 156
Qy      571 ATCCGGACATTACCGCTGCGCAGCTAGCGCCCTC-----CCACACCGGGGGCG 618
Db      156 gProCysSerAsnArgThrSerGlyIleSerMetGlyTyrAspGlnArgHis-glySerP 176
Qy      619 CAATTGACGCCCTCCAGTATTGGCCCTTTCTTCGAGATCTCTATAATGGAAACT 678
Db      176 roLeuProAlaGlyProCysLeuPhe-----GAGCCACCG-----TCACCGACCTACTTGGAT 762
Qy      679 AACCATCCCTTCTTCGAGGATCCCAACGCCCTCACGGGGTGGTGGAGTCCCTTATG 738
Db      185 -----GlyProProLeuAlaGlyAlaProAlaGlyTyrSerProGlyGlyValProSerA 203
Qy      739 TTCTC----- 743
Db      203 laTyrProGluLeuHisAlaAlaLeuAspArgLeuTyrAlaGlnArgProAlaGlyPheG 223
Qy      744 -----TCACCGACCTACTTGGAT 762
Db      223 lyCysGlnGluSerArgHisSerTyrProProAlaLeuGlySerProGlyAlaLeuAlaG 243
Qy      763 GATTGTCAACAGCTGTGCAGACTCTTCAACACCGAGGAGCGAGA-----GAGAATT 816
Db      243 lyAlaGlyValGlyAlaAlaGlyProLeuGluArgArgGlyAlaGlnProGlyArgHis 263
Qy      817 CTATTAGAGCTAGAAAATGTTCTCTGGGCGGCGGAGCGGACCCACCGGTTTCAAAAT 876
Db      263 erValThrGlyTyrGlyAspCysAlaValGlyAlaArgTyrGlnAspGluLeuThr---- 281
Qy      877 GAGATTGCATGGATTCCCTTAACCTGCGCCCGGTTGGGACTACACACGCGCTGAAGT 936
Db      281 ----- 281
Qy      937 AGGAGAGCTTGAATAATCTATCGCAGGCTCTGGTGGCGGTCTCCGGGGCGCTCAAGA 996
Db      282 -----AlaLeuLeuArgLeuThrValGlyThrGlyGlyArgGlu-AlaGlyAlaArgGly 299
Qy      997 CGGCCCACTAATTGGCTAAGTAAAGAGAAGTGTGACGAGGACCGAATGAACCCCTCT 1056
Db      300 GluProSerGlyIleGluPro-----SerGlyLeuGluGluProProGly 314
Qy      1057 GTTTTCTTGAGGCTCTTGGAGCCCTTCAGGCGG 1092
Db      315 ProPheValProGluAlaAlaArgAlaArgMetArg 326

```

```

Qy 817 CTATTAGAGCTAGAAAAATGTTCTCGGGCGGACGGGGCGCCACCGCGGTGCAAAAT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 erValThrGlyTyrClyAspCysAlaValAlaGlyAlaArgTyrGlnAspGluLeuThr 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 877 GAGATTGACATGGGATTTCCCTTAACTCGCCCGGTTGGGACTACAACACGCGTGAAGGT 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 ----- 281
Qy 937 AGGGAGACTTGAATCTATCGGCAGGCTCTGGTGGGGGTCTCCGGGGCGCCTCAAGA 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 -----AlaLeuLeuArgLeuThrValGlyThrGlyArgGlu-AlaGlyAlaArgGly 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 997 CGGCCACTAATTGGCTAAGTAAGAGAAGTGATGCAGGACCGCAATGAACCCCTCT 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 GluProSerGlyIleGluPro-----SerGlyLeuGluGluProProGly 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1057 GTTTTCTTGAGAGGCTCTTGAAGCCCTTCAGCGG 1092
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 PropHeValProGluAlaAlaArgAlaArgMetArg 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-11-124-368A-313
; Sequence 313, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-313

Alignment Scores:
Pred. No.: 0.0123 Length: 538
Score: 136.50 Matches: 83
Percent Similarity: 31.7% Conservative: 23
Best Local Similarity: 24.9% Mismatches: 112
Query Match: 4.8% Indels: 116
DB: 7 Gaps: 12

US-10-723-552-3_copy_585_2156 (1-1572) x US-11-124-368A-313 (1-538)

Qy 301 CTGAATAGCCAGAAACCGGTCCCGAATCTGGCTCTTGA----- 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41 LeuGlyGlyProArgLysSerGlyProArgGlyAlaThrGlyGlyProGlyAspGluPro 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 346 -----GAGAAAAACAAACACATCGGCTGAAAAAGTCAAGCCCTCTCCTCAT 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 LeuGluProAlaArgGluGlnGlySerLeuAspAlaGluArgGlnArgGlySerPhe 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 391 ATCTACCCGAGATTGAG-----GAGCCACCG-----GCTTGGCGG 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 GluAlaProArgTyrGlySerPheProAlaGlyProProProThrArgAlaLeuPro 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 427 GNACCCCAATCTGTTCCCGCA-----CCCTTATCTGGCACAGGTGCCGCG 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 LeuProGlnSerLeuProProAspPheArgLeuGluProThrAlaProAlaLeuSerPro 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 475 AGGGGACCTTTGCC-----CCTCTGGAGCTCCGCGGTG 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 121 ArgSerSerPheAlaSerSerSerAlaSerAspAlaSerLysProSerSerPro-ArgG1 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 511 GAGGGACCTCTGCGAGGACTCGGAGCGGAGGGCGCCACCCCGGAGCGGACAGACGAG 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 ySerLeuLeuLeuAspGlyAlaGlyAlaGlyGlyAla-----GlyGlySerAr 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 571 ATCGCGACATTTACCGTCGCGACGCTAGCGCCCTC-----CCACACCGGGGGCG 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 156 gProCysSerAsnArgThrSerGlyIleSerMetGlyTyrAspGlnArgHis-GlySerP 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 619 CAATTGCGAGCCCTCCAGTATTGCCCTTTTCTTCGAGATCTCTAATTGGAATACT 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 roLeuProAlaGlyProCysLeuPhe----- 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 679 AACCATCCCTTCTTCGAGGATCCCAACGCTCAGGGGTGCTGGAGTCCCTTATG 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 -----GlyProProLeuAlaGlyAlaProAlaGlyTyrSerProGlyValProSerA 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 739 TTCTC----- 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 laTyrProGluLeuHisAlaAlaLeuAspArgLeuTyrAlaGlnArgProAlaGlyPheG 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 744 -----TCACCGCCTTACTTTGGGAT 762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 lyCysGlnGluSerArgHisSerTyrProProAlaLeuGlySerProGlyAlaLeuAlaG 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 763 GATTGTCAACAGCTGCTGCAGACACTCTTCACACCGAGGAGCGAGA-----GAGRATT 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 lyAlaGlyValGlyAlaAlaGlyProLeuGluArgArgGlyAlaGlnProGlyArgHis 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 817 CTATTAGAGGTAGAAAAATGTTCTGGGGCCCGGCGGACCCACCGCGTTGCCAAAT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 erValThrGlyTyrGlyAspCysAlaValAlaGlyAlaArgTyrGlnAspGluLeuThr 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 877 GAGATTGACATGGGATTTCCCTTAACCTCGCCCGGTTGGGACTCAACACGCGTGAAGGT 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 ----- 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 937 AGGGAGAGCTGAAAAATCTATCGCAGGCTCTGGTGGGGGTCTCCGGGGCGCCTCAAGA 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 -----AlaLeuLeuArgLeuThrValGlyThrGlyGlyArgGlu-AlaGlyAlaArgGly 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 997 CGGCCCACTAATTGGTGAAGTAAGAGAGTGAAGTGCAGGAGCGCAATGAACCCCTCT 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 GluProSerGlyIleGluPro-----SerGlyLeuGluGluProProGly 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1057 GTTTTCTTGAGAGGCTCTTGAAGCCCTTCAGCGG 1092
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 PropHeValProGluAlaAlaArgAlaArgMetArg 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-11-059-982-1
; Sequence 1, Application US/11059982
; Publication No. US20050255507A1
; GENERAL INFORMATION:
; APPLICANT: Jenkins, Robert B.
; APPLICANT: Yang, Ping
; APPLICANT: Thibodeau, Steve
; APPLICANT: Wang, Liang
; APPLICANT: Schaid, Daniel
; TITLE OF INVENTION: CYTOGENETICALLY DETERMINED DIAGNOSIS AND
; TITLE OF INVENTION: PROGNOSIS OF PROLIFERATIVE DISORDERS
; FILE REFERENCE: 07039-505001
; CURRENT APPLICATION NUMBER: US/11/059,982
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: US 60/545,573
; PRIOR FILING DATE: 2004-02-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-11-059-982-1

Alignment Scores:

Pred. No.: 0.0131 Length: 1560
Score: 136.50 Matches: 95
Percent Similarity: 29.4% Conservative: 18
Best Local Similarity: 21.0% Mismatches: 156
Query Match: 4.8% Indels: 163
DB: 7 Gaps: 20

US-10-723-552-3_COPY_585_2156 (1-1572) x US-11-059-982-1 (1-1560)

QY 151 CCATCAGGGGACCTTTAATCT-----GAGATTCTCGGCTGTTAAAGCAGTTATT 204
DB 666 ProSerProGlyLeuAlaSerSerProGlyLeuValLeu----- 679
QY 205 TTTGAGCTGACCGGCTCTCATCCGATCAGGAGCCCTATATCTTACGTGCAAGAT 264
DB 680 -----GlyGlnProSerAlaThrProThrAlaIleLeuThrGlnAsp 694
QY 264 ----- 264
DB 695 SerLeuGlnMetPheLeuProGlnGluArgSerGlnGlnProLeuSerAlaGluGlyPro 714
QY 265 -----TTGCAGAGGATCTCCGCCATGCTGGTTAAACCATGG 300
DB 715 HisLeuSerValProAlaSerValIleValSerAlaProProProAlaGlnAspPro--- 733
QY 301 CTGAATAAGCCAGAAAGCCAGGTCGCCGAATCTGGCTCTTGAGAGAGAAAAACACAC 360
DB 734 -----AlaProAlaThrProValAlaGlyAlaGlyLeuGlyProGlnAlaProAsp 751
QY 361 TCGGCTGAAAAGTCAAGCCCTCTCTCATATC----- 393
DB 752 SerGlnAlaSerProAlaProAlaProGlnIleProAlaAlaAlaProLeuLysGlyPro 771
QY 394 -----TACCCCGAGATTGAGAGCCACCGGCTTGG 423
DB 772 GlyProSerSerProSerLeuProHisGlnAlaProLeuGlyAspSerProHisLeu 791
QY 424 CCGGAACCCCAATCTGTTCGCCA-----CCCCCTATCTGGCACAGGTCGCGC 474
DB 792 ProSerProHisProThrArgProProSerArgProProSerArgProGlnSerValSer 811
QY 475 AGGCGACCTTTGCCCCCTCTGAGCTCCCGCGGTGGAGGACCTGCTGCGAGGACT--- 531
DB 812 ArgProSerGluProProLeuHisProCysProProProGlnAlaProProThrLeu 831
QY 532 -----CGAGCCGAGGGGGCGCCACCCCGAGCGGACAGCAGATC 573
DB 832 ProGlyIlePheValIleGlnAsnGlnLeuGlyValProProAlaSerAsnProAla 851
QY 574 GCGACATTACGTCGCGCATGACGGCCCTCC----- 606
DB 852 ProThrAlaPro-----GlyProGlnProProLeuArgProGlnSerGln 867
QY 607 ACACGGGGGCAATTCAGCCCTCCAGTATGGCCCTTTCTCTGCGATGCTC--- 663
DB 868 ProProGluGlyProLeuProProAlaProHisLeuProProSerSerThrSerAla 887
QY 664 -----TATAATTGGAAA 675
DB 888 ValAlaSerSerSerGluThrSerSerArgLeuProAlaProThrProSerAspPheGln 907
QY 676 ACTAACCATCCCCCTTCTCGGAGGATCCCAAGCCCTCAGGGGTGGAGTCCCTT 735
DB 908 LeuGlnPheProPro---SerGlnGlyProHisLysSerProThrProProThrLeu 926
QY 736 ATGTCTCTCACCAGCCT-----ACTTGGGATGATTGT 768
DB 927 HisLeuValProGluProAlaAlaProProProProProProProProProProPhe----- 943
QY 769 CAACAGCTGTGCAGACACTTTTCAACACCGAGGAGCGAGAGAAATTTCTATTAGAGCT 828

DB 944 ---GlnMetValThrThrProPheProAlaLeuProGlnProLysAlaLeuLeuGluArg 962
QY 829 AGAAAAAATGTTCTCTGGG-----GCGAGCGGCGACCCACGCGG 867
DB 963 PheHisGlnValProSerGlyIleIleLeuGlnAsnLysAlaGlyAlaProAlaAla 982
QY 868 TTGCAAAATGAGATTGACATGGGATTTCCCTTAACTCGCCCGGTTGGGACTACAAACAG 927
DB 983 ProGlnThrSerThrSerLeuGly---ProLeuThrSerProAla----- 996
QY 928 GCTGAAGGTAGGAGAGCTTGAAAATCTATGCCAGGCT----- 966
DB 997 -----AlaSerValLeuValSerGlyGlnAlaProSerGlyThrProThrAla 1012
QY 967 -----CTGGTGGCGGCTCTC----- 981
DB 1013 ProSerHisAlaProAlaProAlaProMetAlaAlaThrGlyLeuProProLeuLeuPro 1032
QY 982 -----CGGGCGGCTCAAGACGCGCCACT---AATTTGGCTAAGTAAGAA 1026
DB 1033 AlaGluAsnLysAlaPheAlaSerAsnLeuProThrLeuAsnValAlaLysAla----- 1050
QY 1027 GTGATGCGAGGACCGAATGAACCCCTCTGTTTCTTGTGAGAGCTCTTGGAAAGCCTTC 1086
DB 1051 AlaSerSerGlyProGlyLysProSerGlyLeuGlnTyrgLuserLysLeuSerGlyLeu 1070
QY 1087 AGGCGGTACACCCCTTTTGATCCACCTCAGAGGCC 1122
DB 1071 LysLysProProThrLeuGlnProSerLysGluAla 1082

RESULT 8
US-10-821-234-1050
; Sequence 1050, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes version 1.0
; SEQ ID NO 1050
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1050

Alignment Scores:
Pred. No.: 0.0269 Length: 1410
Score: 132.00 Matches: 84
Percent Similarity: 31.9% Conservative: 58
Best Local Similarity: 18.9% Mismatches: 165
Query Match: 4.6% Indels: 138
DB: 6 Gaps: 16

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-821-234-1050 (1-1410)

QY 346 GAGAAAAACAAACACTCGGCTGAAAAGTCAAGCCCTCTCTCATATCTACCCGAGATT 405
DB 986 GluHisLeuAsnHisSerLysSerValThrProAlaSerThrLeuThrLys----- 1003
QY 406 GAGGAGCCACCGCTGGCGGAAACCCCAATCTGTT----- 441
DB 1004 SerGlyProGlyArgTrpLysThrProAlaAlaIleProAlaThrProValAlaValSer 1023
QY 442 -----CCCCCACCCCTTATCTGCGACAGGGT 468

```
Db      1024  GlnProIleArgThrAspLeuProProProProProProValHisTyAlaGly 1043
Qy      469  CCGCGGAGGGA-----CCCTTGGCCCTCT----- 495
Db      1044  AspPheAspGlyMetSerMetAspLeuProProProProSerAlaAsnGlnIle 1063
Qy      496  CGAGCTCGGCGGTGGAGGACCTCTCGCAGGACTCGGAGCGG----- 540
Db      1064  GlyLeuProSerAlaGlnValAlaAlaGluArgArgLysArgGluGluHisGlnArg 1083
Qy      541  -----AGGGCGCCACCCCGGACCGGACGAC 567
Db      1084  TrpTyrgluLysGluLysAlaArgLeuGluGluArgGluArgLysArgGluGln 1103
Qy      568  GAGATCGGCATATACCGCTCGCAGCTACGGCCCTCCACACCGGGGGCCAATTGCAG 627
Db      1104  GluArgLysLeuGlyGlnMetArgThr-----Gln 1113
Qy      628  CCCCTCCAGTATGCGCCCTTTCTCTCTGCAGATCTCTATAATTGAAACTAACCATCCC 687
Db      1114  SerLeuAsnProAlaProPheSerProLeuThrAlaGlnGlnMetLysProGluLysPro 1133
Qy      688  CCTTCTCGAGGATCCCAACCGCTCACGGGTTGGTGAGTCCCTTATGTCTCTCAC 747
Db      1134  SerThrLeuGlnArgProGlnGluThr-----ValIleArgGluLeuGlnProGlnGln 1151
Qy      748  CAGCCTACTGGGATGATGTCAACAGCTGTGCAGACACTTTCACAAACGAGGAGGA 807
Db      1152  GlnProArg-----ThrIleGluArgArg 1159
Qy      808  GAGAGATTCTATTAGCGCTAGAAAAATGTTCTCTGGGCGGACGGCGGACCGCGG 867
Db      1160  AspLeuGlnTyrlleThrValSerLys----- 1168
Qy      868  TTGCAAAATGAGATTGACATGGGATTTCCCTTAACTCGCCCGGTGGGACTACAACACG 927
Db      1169  -----GluGluLeuSerSerGlyThrValCysProProAspProTrp-----Lys 1183
Qy      928  GCTGAGGTGAGGAGAGCTTGAAATCTATCGCCAG-----GCTCTGGTG 972
Db      1184  ArgAlaLysGluLysLeuGluLysGlnGlnMetHisIleValAspMetLeuSer 1203
Qy      973  CGGGTCTCCGGGGCGCTCAAGCGGCCACTAATTGGCT-----AAGTA 1020
Db      1204  LysGluIleGlnGluLeuGlnSerLysProAspArgSerAlaGluLysSerAspArgLeu 1223
Qy      1021  AGAGAAGTGATG----- 1032
Db      1224  ArgLysLeuMetLeuGluTrpGlnPheGlnLysArgLeuGlnGlnSerLysGlnLysAsp 1243
Qy      1033  -----CAGGACCGNATGAACCCCTCTGTTTTTCTTGAGAGGCTCTTG 1077
Db      1244  GluAspAspGluGluGluAspAspValAspThrMetLeuIleMetGlnArgLeu 1263
Qy      1078  GAAGCTTCAGCGGTACACCCCTTTGATCCACCTCAGAGGCCCAAAAGCCCTCAGTG 1137
Db      1264  GluAlaGluArg-----AlaArgThrAlaMetPro 1274
Qy      1138  GCTTTGGCTTTATTAGGACAGTCAGCCTTGGATATTAGAAAGAGCTTCAGAGCTGGAA 1197
Db      1275  AlalleSerValLeuAspLeuGlnAspGluGluArgArgGlnGlnGlnGlu 1294
Qy      1198  GGGTTACAGGCGGTGAGTTACGTGATCTAGTGAAGGAGGAGGAGAAAGTATTACAA 1257
Db      1295  GluMetArgLysArgGluAlaGluAspArgAlaArgGlnGluGlu-----Arg 1311
Qy      1258  AGGGAGACAGAGAGAGGAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
Db      1312  ArgArgGlnGluGluGluArgThrLysArgAspAlaGluGluLysArgArgGlnGlu 1331
Qy      1318  -----AGACCTAATAACGGCAAGAGAAATTTGACTAAG 1353
```

```
Db      1332  GlyTyrtyrSerArgLeuGluAlaGluArgArgGlnHisAspGluAlaAlaArgArg 1351
Qy      1354  ATCTTGCTGCAGTGGTTGAGGAGGAAAAGCAATACGGAAGAGAGAGAGATTTAGGAA 1413
Db      1352  LeuLeuGluProGluAlaProGlyLeuCysArgProProLeuProAArgAspTyrgluPro 1371
Qy      1414  ATTAGGTGAGGCCCT 1428
Db      1372  ProSerProSerPro 1376

RESULT 9
US-11-124-367A-444
; Sequence 444, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 444
; LENGTH: 2760
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-124-367A-444

Alignment Scores:
Pred. No.:      0.0419      Length:      2760
Score:          129.50      Matches:      71
Percent Similarity: 42.3%      Conservative: 72
Best Local Similarity: 21.0%      Mismatches: 119
Query Match:      4.5%      Indels:      76
DB:              7        Gaps:      16

US-10-723-552-3_COPY_585_2156 (1-1572) x US-11-124-367A-444 (1-2760)
Qy      712  CTCACGGGGTGTGTGAGTCCCTTATGTTCTCTCACAGCCTACTTGGGATGATTGTCAA 771
Db      247  IleThrAlaLeuVal-----ValTrpLeuProAlaAsnProThr-----Glu 260
Qy      772  CAGCTGTGTGCAGACACTCTTCACAAACCGAGGAGGAGAGAAATTTCTATTAGAGGCTAGA 831
Db      261  LysIleValArgValLeuPheProGlyAsnAlaProGlnAsnLysIleLeuGluGlyLeu 280
Qy      832  AAAAAATGTTCTCGGGCGGAC-----GGGCGACCCCGCGGTTCGAAAAATGAGATTGAC 885
Db      281  GluLysLeuArgHisLeuAspPheLeuArgTyrgluAlaThrGlnLysAspLeuAla 300
Qy      886  ATGGGATTTCCCTTAACCT-----CGCCCCGGTTGGGACTACACACCGCTGAAGGTAGG 939
Db      301  SerGlyAlaValProThrAsnLeuLysProSerLysIleLysGlnArgAlaAspSerLys 320
Qy      940  GAGAGCTTGAATACTATCGCCAGGCTCTGGTGGCGGCTC----- 981
Db      321  GluSerLeuLysAlaThrThrLysThrAlaValSerLysLeuAlaLysArgGluGluVal 340
Qy      982  -----CGGGCGCTCAAGACCGCCCACTAATTTGGCTAAG-----GTA 1020
Db      341  ValGluGluGlyAlaLysGluAlaArgSerGluLeuAlaLysGluLeuAlaLysThrGlu 360
Qy      1021  AGAGAAGTGTGAGGAGGACCGAATGAACCCCTCTGTTTTTCTTGAGAGGCTCTTGAA 1080
Db      361  LysLysAlaLysGluSerSerGluLysProPro-----GluLysProAlaLys 376
```


QY	1081	GCCTTCAGCGCGGTACACCCCTTTTGATGCCACCTCAGAGGCCCAAAAGACCTCAGTGGCT	1140
Db	377	ProGluArgValLysThrGluSerSerGluAlaLeuLysAlaGluLysArgLysLeuIle	396
QY	1141	TTGGCCCTTTATAGACAGCTCAGCCCTCGATATTAGAAAGAGCTTCAGACACTGGAGGG	1200
Db	397	LysAspLysValGlyLysLysHisLeu-----LysGluLysIleSerLysLeuGluGlu	414
QY	1201	TTACAGGAGCGCTGAGTTACGTGATCTAGTGAAGAGGCAGACGAAGATATATTACAAAAGG	1260
Db	415	LysLysAspLysGluLysLysGluIleLysLysGluArgLysGluLeu-----Lys	431
QY	1261	GAGACAGAGNAGAAAGGGAACAAGAAAGACAGACAGAAAGAGGAAAGCGGAGGAAAGA	1320
Db	432	LysAspGluGlyArgLysGluGluLysLysAspAlaLysLysGluGluLys---ArgLys	450
QY	1321	CGTAATAAACGGCAAGAGNAGAAATTCACATAAG-----Lys	1353
Db	451	AspThrLysProGluLeuLysLysLysLysLysProAspLysLysProPheThrProGlu	470
QY	1354	-----AFTCTTGGCTGCAGTGGTTGAGGCGAAACCAATACGAAAGAGAGAGA	1401
Db	471	ValArgLysThrLeuTyrLysAlaLysValProGlyArgValLysIleAspArgSerArg	490
QY	1402	GATTTTAGG-----AAATTAGCTCAGCCCTCAGACAGTCAAGG	1440
Db	491	AlaIleArgGlyGluLysGluLeuSerSerGluProGlnThrProAlaGlnLysGly	510
QY	1441	AACCTG-----GGCAATAGG-----ACCCCACTCGAC	1467
Db	511	ThrValProLeuProThrIleSerGlyHisArgGluLeuValLeuSerSerProGluAsp	530
QY	1468	-----AAGGACCAATGTGCATATTCTGAAGAAAGA	1497
Db	531	LeuThrGlnAspPheGluGluMetLysArgGluGluArgAlaLeuLeuAlaGluGlnArg	550
QY	1498	GGACACTGGCGAGGAACCTGCCCA-----AGAGGGGAAACAAAGACCAAGGA	1546
Db	551	AspThrGlyLeuGlyAspLysProPheProLeuAspThrAlaGlnGlyGly	567

```

RESULT 10
US-11-124-367A-442
; Sequence 442, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 442
; LENGTH: 2803
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-367A-442

```

Alignment Scores:	
Pred. No.:	0.0419
Score:	129.50
Length:	2803
Percent Similarity:	42.3%
Matches:	71
Best Local Similarity:	21.0%
Conservative:	72
Query Match:	119
Mismatches:	119
Indels:	76

[illegible]

RESULT 11


```
QY 832 AAAAATGTTCTGGGGCCGAC-----GGGGGACCCACGCGGTTGCAAAATGAGATTGAC 885
   : : : : :
Db 505 GluLysLeuArgHisLeuAspPheLeuArgTyrProValAlaThrGlnLysAspLeuAla 524
   : : : : :
QY 886 ATGGGATTTCCCTTAAC-----CGCCCGGTTGGGACTACACACGCGCTGAAGTAGG 939
   : : : : :
Db 525 SerGlyAlaValProThrAsnLeuLysProSerLysIleLysGlnArgAlaAspSerLys 544
   : : : : :
QY 940 GAGAGCTTGAATAATCTATCGCCAGGCTCTGGTGGCGGCTC----- 981
   : : : : :
Db 545 GluSerLeuLysAlaThrThyLysThrAlaValSerLysLeuAlaLysArgGluVal 564
   : : : : :
QY 982 -----CGGGCGGCTCAAGCGGCCCACTAATTGGCTAAG-----GTA 1020
   : : : : :
Db 565 ValGluGluGlyAlaLysGluAlaArgSerGluLeuAlaLysGluLeuAlaLysThrGlu 584
   : : : : :
QY 1021 AGAAGATGATGACGGGACCGAATGAACCCCTCTGTTCTTTCGAGAGCTCTGGAA 1080
   : : : : :
Db 585 LysLysAlaLysGluSerSerGluLysPro-----GluLysProAlaLys 600
   : : : : :
QY 1081 GCCTTCAGGCGGTACACCCCTTTTGATCCACCTCAGAGGCCCAAAAGCCTCAGTGCT 1140
   : : : : :
Db 601 ProGluArgValLysThrGluSerSerGluAlaLysAlaGluLysArgLysLeuLe 620
   : : : : :
QY 1141 TTGGCTTTATAGACAGCTCAGCTCGATATAGAAAGAGCTTCAGAGACTCGAAGGG 1200
   : : : : :
Db 621 LysAspLysValGlyLysHisLeu-----LysGluLysIleSerLysLeuGluGlu 638
   : : : : :
QY 1201 TTACAGGAGCTGAGTTACGTATCTAGTGAAGGAGCAGAGAAAGTATATACAAAAGG 1260
   : : : : :
Db 639 LysLysAspLysGluLysGluLysIleLysLysGluLysGluLysGluLysGluLys 655
   : : : : :
QY 1261 GAGACAGAAAGAAAGGGAACAAAGAAAGAGAGAGAAAGAGAGAGAGAGAGAGAGAG 1320
   : : : : :
Db 656 LysAspGluGlyArgLysGluLysGluLysLysAspAlaLysLysGluLysGluLys 674
   : : : : :
QY 1321 CGTAATAAAGCGCAAGAGAAAGTTCACCTAAG----- 1353
   : : : : :
Db 675 AspThrLysProGluLeuLysLysIleSerLysProAspLeuLysProPheThrProGlu 694
   : : : : :
QY 1354 -----ATCTTGGCTGCAGTGTGTTGAAGGGAACCAATACGGAAGAGAGAGA 1401
   : : : : :
Db 695 ValArgLysThrLeuTyrLysAlaLysValProGlyArgValLysIleAspArgSerArg 714
   : : : : :
QY 1402 GATTTTAGG-----AAATTAGCTCAGCGCCCTAGACAGCTCAGGG 1440
   : : : : :
Db 715 AlaIleArgGlyGluLysGluLeuSerSerGluProGlnThrProProAlaGlnLysGly 734
   : : : : :
QY 1441 AACCTG-----GGCAATAGG-----ACCCCACTCGAC 1467
   : : : : :
Db 735 ThrValProLeuProThrLysSerGlyHisArgGluLeuValLeuSerSerProGluAsp 754
   : : : : :
QY 1468 -----AAGACCAATGTGCATATTGTAAGAAAGA 1497
   : : : : :
Db 755 LeuThrGlnAspPheGluGluMetLysArgGluGluArgAlaLeuLeuAlaGlnLysArg 774
   : : : : :
QY 1498 GGACACTGGGCAAGCAAGTGGCCCA-----AGAAGGGAACCAAGCACAAGA 1546
   : : : : :
Db 775 -AspThrGlyLeuGlyAspLysProPheProLeuAspThrAlaGluGly 791
   : : : : :
```

RESULT 13

```
US-11-124-367A-441
; Sequence 441, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
```

```
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 441
; LENGTH: 3027
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-124-367A-441
```

Alignment Scores:

Pred. No.:	0.0421	Length:	3027
Score:	129.50	Matches:	71
Percent Similarity:	42.3%	Conservative:	72
Best Local Similarity:	21.0%	Mismatches:	119
Query Match:	4.5%	Indels:	76
DB:	7	Gaps:	16

US-10-723-552-3_COPY_585_2156 (1-1572) x US-11-124-367A-441 (1-3027)

```
QY 712 CTCACGGGTTGGTGGAGTCCCTTATGTTCTCTCACCAGCCTACTTGGGATGATTGTCAA 771
   : : : : :
Db 471 IleThrAlaLeuVal-----ValTrpLeuProAlaAsnProThr-----Glu 484
   : : : : :
QY 772 CAGCTGCTGCAGACACTTTCACAAACGAGGAGCGAGAGAGAAATTCATTATAGAGCTAGA 831
   : : : : :
Db 485 LysIleValArgValLeuPheProGlyAsnAlaProGlnAsnLysIleLeuGluGlyLeu 504
   : : : : :
QY 832 AAAAATGTTCTCTGGGGCCGAC-----GGGGGACCCACGCGGTTGCAAAATGAGATTGAC 885
   : : : : :
Db 505 GluLysLeuArgHisLeuAspPheLeuArgTyrProValAlaThrGlnLysAspLeuAla 524
   : : : : :
QY 886 ATGGGATTTCCCTTAAC-----CGCCCGGTTGGGACTACAAACGCGCTGAAGTAGG 939
   : : : : :
Db 525 SerGlyAlaValProThrAsnLeuLysProSerLysIleLysGlnArgAlaAspSerLys 544
   : : : : :
QY 940 GAGAGCTTGAATAATCTATCGCCAGGCTCTGGTGGCGGCTC----- 981
   : : : : :
Db 545 GluSerLeuLysAlaThrThyLysThrAlaValSerLysLeuAlaLysArgGluVal 564
   : : : : :
QY 982 -----CGGGCGGCTCAAGAGCGGCCCACTAATTGGCTAAG-----GTA 1020
   : : : : :
Db 565 ValGluGluGlyAlaLysGluAlaArgSerGluLeuAlaLysGluLeuAlaLysThrGlu 584
   : : : : :
QY 1021 AGAAGATGATGACGGGACCGAATGAACCCCTCTCTGTTTCTTTCGAGAGGCTCTTGAA 1080
   : : : : :
Db 585 LysLysAlaLysGluSerSerGluLysPro-----GluLysProAlaLys 600
   : : : : :
QY 1081 GCCTTCAGGCGGTACACCCCTTTTGATCCACCTCAGAGGCCCAAAAGCCTCAGTGCT 1140
   : : : : :
Db 601 ProGluArgValLysThrGluSerSerGluAlaLysAlaGluLysArgLysLeuLe 620
   : : : : :
QY 1141 TTGGCTTTATAGACAGCTCAGCTCGATATAGAAAGAGCTTCAGAGACTCGAAGGG 1200
   : : : : :
Db 621 LysAspLysValGlyLysHisLeu-----LysGluLysIleSerLysLeuGluGlu 638
   : : : : :
QY 1201 TTACAGGAGCTGAGTTACGTATCTAGTGAAGGAGCAGAGAAAGTATATACAAAAGG 1260
   : : : : :
Db 639 LysLysAspLysGluLysGluLysIleLysLysGluLysGluLysGluLysGluLys 655
   : : : : :
QY 1261 GAGACAGAAAGAAAGGGAACAAAGAAAGAGAGAGAAAGAGAGAGAGAGAGAGAGAG 1320
   : : : : :
Db 656 LysAspGluGlyArgLysGluLysGluLysLysAspAlaLysLysGluLysGluLys 674
   : : : : :
QY 1321 CGTAATAAAGCGCAAGAGAAAGTTCACCTAAG----- 1353
   : : : : :
Db 675 AspThrLysProGluLeuLysLysIleSerLysProAspLeuLysProPheThrProGlu 694
   : : : : :
QY 1354 -----ATCTTGGCTGCAGTGTGTTGAAGGGAACCAATACGGAAGAGAGAGA 1401
   : : : : :
Db 695 ValArgLysThrLeuTyrLysAlaLysValProGlyArgValLysIleAspArgSerArg 714
   : : : : :
QY 1402 GATTTTAGG-----AAATTAGCTCAGCGCCCTAGACAGCTCAGGG 1440
   : : : : :
Db 715 AlaIleArgGlyGluLysGluLeuSerSerGluProGlnThrProProAlaGlnLysGly 734
   : : : : :
QY 1441 AACCTG-----GGCAATAGG-----ACCCCACTCGAC 1467
   : : : : :
Db 735 ThrValProLeuProThrLysSerGlyHisArgGluLeuValLeuSerSerProGluAsp 754
   : : : : :
QY 1468 -----AAGACCAATGTGCATATTGTAAGAAAGA 1497
   : : : : :
Db 755 LeuThrGlnAspPheGluGluMetLysArgGluGluArgAlaLeuLeuAlaGlnLysArg 774
   : : : : :
QY 1498 GGACACTGGGCAAGCAAGTGGCCCA-----AGAAGGGAACCAAGCACAAGA 1546
   : : : : :
Db 775 -AspThrGlyLeuGlyAspLysProPheProLeuAspThrAlaGluGly 791
   : : : : :
```

```
Db      695 ValArgLysThrLeuTyrlsAlaLysValProGlyArgValLysIleAspArgSerArg 714
      :::      |||      |||      |||      |||      |||      |||      |||
Qy      1402 GATTTTAGG-----AAATTAGGTGAGGCCCTAGACAGTCAGG 1440
      |||      |||      |||      |||      |||      |||      |||
Db      715 AlaIleArgGlyGluLysGluLeuSerGluProGlnThrProAlaGlnLysGly 734
      |||      |||      |||      |||      |||      |||      |||
Qy      1441 AACCTG-----GGCAATAGG-----ACCCACTCGAC 1467
      |||      |||      |||      |||      |||      |||      |||
Db      735 ThrValProLeuProThrIleSerGlyHisArgGluLeuValLeuSerSerProGluAsp 754
      |||      |||      |||      |||      |||      |||      |||
Qy      1468 -----AAGGACCAATGTCATATTGTAAGAAAGA 1497
      |||      |||      |||      |||      |||      |||      |||
Db      755 LeuThrGlnAspPheGluGluMetLysArgGluGluArgAlaLeuLeuAlaGluGlnArg 774
      |||      |||      |||      |||      |||      |||      |||
Qy      1498 GGACACTGGCAAGAACTGCCCA---AGAAAGGAAACAAAGGACCAAGGA 1546
      |||      |||      |||      |||      |||      |||      |||
Db      775 -AspThrGlyLeuGlyAspLysProPheProLeuAspThrAlaGluGluGly 791
      |||      |||      |||      |||      |||      |||      |||

RESULT 14
US-11-182-016-22
; Sequence 22, Application US/11182016
; Publication No. US20060019294A1
; GENERAL INFORMATION:
; APPLICANT: SUGEN, INC.
; TITLE OF INVENTION: TYROSINE KINASE SUBSTRATE (TKS) PROTEINS
; FILE REFERENCE: 038602/0102
; CURRENT APPLICATION NUMBER: US/11/182,016
; CURRENT FILING DATE: 2005-07-15
; PRIOR APPLICATION NUMBER: US/09/958,359
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 1744
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Grub
US-11-182-016-22

Alignment Scores:
Pred. No.: 0,042      Length: 1744
Score: 129.00      Matches: 124
Percent Similarity: 33.0%      Conservative: 63
Best Local Similarity: 21.9%      Mismatches: 162
Query Match: 4.5%      Indels: 217
DB: 7      Gaps: 30

US-10-723-552-3_COPY_585_2156 (1-1572) x US-11-182-016-22 (1-1744)
Qy      213 TGGACCCGGCTCTCATCCCGATCAGGAGCCCTATATCTTACGTG---GCAAGATTGGC 269
      |||      |||      |||      |||      |||      |||      |||
Db      670 TrpSerGlnLeuArg-----AlaLeuPheLeuIleValLeuSerThrTrpGly 685
      |||      |||      |||      |||      |||      |||      |||
Qy      270 AGAGATCTCGCCATGGTTAAACCATGGCTGAATAAGCCAAAGACCGAGTCCCG 329
      |||      |||      |||      |||      |||      |||      |||
Db      686 GlnGlyAlaMetGly-TrpGlyTrpProTrp-----ProAr 697
      |||      |||      |||      |||      |||      |||      |||
Qy      330 AATTCTGGCTCTTGAGAGAAACAAACACTCGGCTGAAAAGTCAAGCCCTCTCTCA 389
      |||      |||      |||      |||      |||      |||      |||
Db      697 9-----AlaAlaGlyGluGlnGlnGlyAlaGlyArgGlyAlaAlaGlyProPr 715
      |||      |||      |||      |||      |||      |||      |||
Qy      390 TATCTACCCCGAGATTGAGGAG-----CCACCGGCTTGGCCGGAACCCCAATCT-- 438
      |||      |||      |||      |||      |||      |||      |||
Db      715 oLeuGlyProSerThrAlaGlyThrValCysProProSerSerValAlaProTrpAlaAr 735
      |||      |||      |||      |||      |||      |||      |||
Qy      439 -----GTTCCCGCCACCCCTTATCTGGCACAGGCTGCGCGAGGGG 479
      |||      |||      |||      |||      |||      |||      |||
Db      735 gLeuLeuSerGlnValLeuLeuAspProProLeuLeuLeuAlaAlaGlyLeuAlaAla 755
      |||      |||      |||      |||      |||      |||      |||
Qy      480 ACCCTTTGCCCT-----CCTGGAGCTCCGCGCGGTGA 512
      |||      |||      |||      |||      |||      |||      |||
```

```
Qy 1314 GGAAGACGTAAATA-----AACGCAAGAGAGAAATTGACTAAGATCTTTGGCTGC 1364
Db 1073 rgIleSerValLeuLysLeuAlaGlnGlyArgAsn-----GluCysTrpArgG 1089
Qy 1365 AGTCG-----TTGAAGGGAA 1379
Db 1089 InTrpArgCysLeuThrLysValGlnGlnAspGlyGlnProGlnCysLeuLysLeuIleArgProG 1109
Qy 1380 AAGCAATACGAAAGAGAGAGAGATTTTAGGAAATTA-----GGTCAGGCCCTAGACA 1433
Db 1109 luGlnValGluValMetLysSerCysValGlnArgValSerGlyGlyTrpGlyLeuLeuA 1129
Qy 1434 GTCAGGGAACCTGGCAATAGGACCCCACTGCAAGAGCAATGTGCATATTGTAAGA 1493
Db 1129 rgArgAlaArgTrpGlyAsnGlyGlnGlnSerProThrLeuValHisProSerArgG 1149
Qy 1494 AAGAGGAC 1501
Db 1149 InAspGlu 1151

RESULT 15
US-10-982-545-14
; Sequence 14, Application US/10982545
; Publication No. US20050244890A1
; GENERAL INFORMATION:
; APPLICANT: Davies, Huw Alun
; APPLICANT: McGuire, James
; APPLICANT: Simonsen, Anja Hviid
; APPLICANT: Blennow, Kaj
; APPLICANT: Podust, Vladimir
; APPLICANT: CIPHERGEN Biosystems, Inc.
; TITLE OF INVENTION: Biomarkers for Alzheimer's Disease
; FILE REFERENCE: 016866-011550US
; CURRENT APPLICATION NUMBER: US/10/982,545
; CURRENT FILING DATE: 2004-11-06
; PRIOR APPLICATION NUMBER: US 60/519,360
; PRIOR FILING DATE: 2003-11-07
; PRIOR APPLICATION NUMBER: US 60/526,753
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US 60/546,423
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US 60/547,250
; PRIOR FILING DATE: 2004-02-23
; PRIOR APPLICATION NUMBER: US 60/558,896
; PRIOR FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US 60/572,617
; PRIOR FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: US 60/586,503
; PRIOR FILING DATE: 2004-07-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: nerve growth factor inducible Neurosecretory
; OTHER INFORMATION: protein vaccinia virus growth factor (VGF)
; OTHER INFORMATION: precursor
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(22)
; OTHER INFORMATION: signal peptide
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (23)..(615)
; OTHER INFORMATION: nerve growth factor inducible Neurosecretory
; OTHER INFORMATION: protein vaccinia virus growth factor (VGF)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (373)..(417)
; OTHER INFORMATION: biomarker peptide 4808 Da, 4812 Da (Q10),
```

```
; OTHER INFORMATION: processed fragment of nerve growth factor
; OTHER INFORMATION: inducible Neurosecretory protein vaccinia virus
; OTHER INFORMATION: growth factor (VGF)
US-10-982-545-14

Alignment Scores:
Pred. No.: 0.045 Length: 615
Score: 128.50 Matches: 100
Percent Similarity: 32.8% Conservative: 44
Best Local Similarity: 22.8% Mismatches: 176
Query Match: 4.5% Indels: 119
DB: 6 Gaps: 19

US-10-723-552-3_COPY_585_2156 (1-1572) x US-10-982-545-14 (1-615)
Qy 412 CCACCGCTGGCGGGAACCCCAATCTGTCCCCACCCCTTATCTGGCACAGGGTCC 471
Db 24 ProProGlyArgProGluAlaGln-----ProProLeuSerSerGluHisLys 40
Qy 472 GCGAGGGGACCTTTTGCCCTCTCTGGAGCTCCGCGGTGGAGGACCTGTGCGAGGACT 531
Db 41 -----GluProValAlaGlyAspAlaValProGlyProLysAspGlySer 55
Qy 532 -----CGAGCGCGGAGGCGGCCACCCCGGAGCGGACAGACGAG----- 570
Db 56 AlaProGluValArgGlyAlaArgAsnSerGluProGlnAepGluGlyGluLeuPheGln 75
Qy 571 -----ATCGCGACATTACCCCTG-----CCGACGTACGGC 600
Db 76 GlyValAspProArgAlaLeuAlaValLeuLeuGlnAlaLeuAepArgProAlaSer 95
Qy 601 CTTCCACACCGGGGGCCCAATTGCGAG-----CCCTCCAGTATTGGCCCTTTCTCT 627
Db 96 ProProAlaProSerGlySerGlnGlnGlyProGluGluGluAlaGluAlaLeuLeu 115
Qy 628 -----CCCTCCAGTATTGGCCCTTTCTCTCTCCAGTATTGGCCCTTTCTCTCT 654
Db 116 ThrGluThrValArgSerGlnThrHisSerLeuProAlaProGluSerProGluProAla 135
Qy 655 GCAGATCTCTAATTTGGAAACTAACCATCCCTCTTCTCGGAGGATCCC----- 705
Db 136 AlaProArgProGlnThrProGluAsnGlyProGluAlaSerAspProSerGluGlu 155
Qy 706 ---CAACGCTCACGGGTTGGTGAGTCCCTTATGTTCTCTCACACGCTACTTGGGAT 762
Db 156 LeuGluAlaLeuAlaSerLeuLeuGlnGluLeuArg---AspPheSerProSerSerAla 174
Qy 763 GATTGTCAACAGCTGCTGCAGACACTCTTCCACACCGGAGGCGAGAGAAATCTTATTA 822
Db 175 LysArgGlnGlnGluThrAlaAlaAala--GluThrGluThrArgThrHisThrLeuThr 193
Qy 823 GAGGCTAGAAAAAATGTTCTCTGGGCGCGAC----- 852
Db 194 ArgValAsnLeuGluSerProGlyProGluArgValTrpArgAlaSerTrpGlyGluPhe 213
Qy 853 -----GGCGACCCCGCTTGCAGAAAT 876
Db 214 GlnAlaArgValProGluArgAlaProLeuProProProAlaProSerGlnPheGlnAla 233
Qy 877 GAGATTGACATGGGATTTCCCTTAACCTCGCCCGGTTGGGACTACACACGCTGAAGGT 936
Db 234 ArgMetProAspSerGlyProLeuProGlu-----ThrHisLysPheGlyGluGly 250
Qy 937 AGGAGAGCTTGAATAATCTATCGCCAGGCTCTGTGCGGGGTCTCCGGGCGGCC----- 990
Db 251 ValSerSerProLysThrHisLeuGlyGluAlaLeuAlaProLeuSerLysAlaTyrGln 270
Qy 991 -----TCAAGACGCGCCCACTAAATTTGGCTAAGGTAAGGAA 1026
Db 271 GlyValAlaAlaProPheProLysAlaArgArgPro-----GluSer 284
Qy 1027 GTGATCGAGGACCGAATGAACCCCTCTGTGTTTTTTCTTGAGAGGCTCTTG----- 1077
Db 1027 GTGATCGAGGACCGAATGAACCCCTCTGTGTTTTTTCTTGAGAGGCTCTTG----- 1077
```

```
Db      285 AlaLeuLeuGlyGlySerGluAlaGlyGluArgLeuLeuGlnGlnGlyLeuAlaGlnVal 304
Qy      1078 GAAGCCTTCAGCGGTACACCCCTTTTCATCCACCTCAGAGGCCCAAAAAGCCTCAGTG 1137
Db      305 GluAlaGlyArgArgGlnAlaGluAlaThrArgGlnAlaAlaAlaGlnGluGluArgLeu 324
Qy      1138 GCT---TTGGCCTTTATAGGACAGTCAGCCTTGGATATTAGAAAGAGCTTCAG----- 1188
Db      325 AlaAspLeuAla-----SerAspLeuLeuGlnTyrLeuLeuGlnGlyGly 340
Qy      1189 ---AGACTGGAGGTTTACAGGAGGCTCAGTTACCTGATCTAGTCAAGGAGGCAGAGAAA 1245
Db      341 AlaArgGlnArgGlyLeuGlyArgGlyLeuGlnGluAlaAlaGluGluArgGluSer 360
Qy      1246 GTATATTACAAGAGGAGACAGACAGAAAGGGAACAAAGAAAGAGAGAGAAAGAGAG 1305
Db      361 AlaArgGluGluGluGluAlaGluGlnGluArgArgGlyGlyGluGluArgValGlyGlu 380
Qy      1306 GAAAGGGAGGAAAGACGTAATAACGGCAAGAG-----AAGAAT 1344
Db      381 GluAspGluGluAlaAlaGluAlaGluAlaGluAlaGluAlaGluArgAlaArgGln 400
Qy      1345 TTGACTAAGATCTTGGCTGCAGTGTTCGAGGGAAGCAATACGGAAGAGAGAGAGAT 1404
Db      401 AsnAlaLeuLeuPheAlaGluGluGluAspGlyGluAlaGlyAlaGluAspLysArgSer 420
Qy      1405 -----TTTAGGAAAATTAGTCAGGCCCTAGACAGCTCAGGG 1440
Db      421 GlnGluGluThrProGlyHisArgArgLysGluAlaGluGlyThrGluGluGlyGly 439
```

Search completed: February 14, 2006, 17:20:48
Job time : 48.8924 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 14:51:21 ; Search time 20.0582 Seconds
(without alignments)
3295.451 Million cell updates/sec

Title: US-10-723-552-3_COPY_2307_5741
Perfect score: 6183
Sequence: 1 ATGGGTGCCAGGCGCAACA.....CTGTCATTAACCTCTCAGAC 3435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10723552/runat 14022006 125143 12937/app query.fasta_1
-DB-PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USBR=US10723552 @CGN 1 1 107 @runat 14022006 125143 12937 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80:*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4135.5	66.9	1165	1	GNLJGL HIV-1 retropepsin
2	3836	62.0	1196	1	GNMVGV HIV-1 retropepsin
3	3818.5	61.8	1204	2	S35475 pol polyprotein -
4	3809.5	61.6	1204	2	S70393 pol polyprotein -
5	3802.5	61.5	1189	1	GNMVW7 HIV-1 retropepsin
6	3797.5	61.4	1199	1	GNMVW1 HIV-1 retropepsin
7	3778	61.1	1196	1	GNMVWV HIV-1 retropepsin
8	3696	59.8	1784	2	T10532 gag-pol polyprotei
9	3536.5	57.2	1046	1	GNMVCE pol polyprotein -
10	2639.5	42.7	843	1	GNVWK pol polyprotein -
11	1725.5	27.9	476	2	S04842 pol polyprotein -
12	1681	27.2	581	2	A42743 pol polyprotein -
13	1572	25.4	1751	2	T09394 gag-pro-pol polypr
14	1147	18.6	409	2	S12588 pol polyprotein -

15	1142.5	18.5	602	2	B46312 pol polyprotein -
16	1094	17.7	559	2	A46311 pol polyprotein -
17	1084	17.5	390	2	A26621 retrovirus-related
18	1029.5	16.7	294	2	A05071 pol polyprotein -
19	1023.5	16.6	848	4	A44282 retrovirus-related
20	984.5	15.6	282	2	A26103 pol polyprotein -
21	905.5	14.6	473	2	A03959 pol polyprotein -
22	811	13.1	1157	1	GNLJLK pol polyprotein -
23	801	13.0	1161	2	S18738 pol protein - simi
24	692	11.2	886	1	GNLJSP pol polyprotein -
25	678.5	11.0	958	2	S15566 pol protein - simi
26	662.5	10.7	1182	2	T29097 pro-pol-dutPase po
27	639	10.3	982	1	GNLJH2 pol polyprotein -
28	604	9.8	1053	1	GNLJBT HIV-1 retropepsin
29	601	9.7	1146	1	GNLJ22 HIV-1 retropepsin
30	601	9.7	1145	1	GNLJEM HIV-1 retropepsin
31	599.5	9.7	1145	1	GNLJEV HIV-1 retropepsin
32	587	9.5	867	1	GNLJMP pol polyprotein (c
33	578	9.3	867	1	GNLJSA pol polyprotein -
34	573.5	9.3	896	1	GNLJGH pol polyprotein -
35	570	9.2	697	2	A26132 gag-abl-pol polypr
36	567.5	9.2	896	1	GNLJCN pol polyprotein ho
37	563.5	9.1	1295	2	S60179 pol polyprotein ho
38	561.5	9.1	1611	2	G84493 probable retroelem
39	561.5	9.1	2272	2	T18572 gag, pol and env p
40	552.5	8.9	399	4	A45689 retrovirus-related
41	543.5	8.8	852	1	GNLJGA pol polyprotein -
42	541.5	8.8	852	2	S29358 pol protein - bovi
43	538.5	8.7	888	1	GNLJHD pol polyprotein -
44	538	8.7	843	2	S33123 pol polyprotein -
45	531.5	8.6	1313	2	T29193 hypothetical prote

ALIGNMENTS

RESULT 1

GNLJGL
HIV-1 retropepsin (EC 3.4.23.16) - gibbon ape leukemia virus
N;Contains: nuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polymerase
C;Species: gibbon ape leukemia virus
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: B32595
R;Delassus, S.; Sonigo, P.; Wain-Hobson, S.
Virology 173, 205-213, 1989
A;Title: Genetic organization of gibbon ape leukemia virus.
A;Reference number: A32595; MUID:90051069; PMID:2683360
A;Accession: B32595
A;Molecule type: genomic RNA
A;Residues: 1-1165
A;Cross-references: UNIPROT:P21414; UNIPARC:UPI0000131EDE; GB:M26927; NID:g332610; PIDN
C;Comment: The pol polyprotein contains reverse transcriptase and possibly an endonucle
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase; polyp
F;3-102/Product: retropepsin #status predicted <RTP>
F;27/Active site: Asp (shared with dimeric partner) #status predicted

Alignment Scores:

Pred. No.: 1.12e-268 Length: 1165
Score: 4135.50 Matches: 777
Percent Similarity: 79.7% Conservative: 128
Best Local Similarity: 68.4% Mismatches: 200
Query Match: 66.9% Indels: 31
DB: 1 Gaps: 7

US-10-723-552-3_COPY_2307_5741 (1-3435) x GNLJGL (1-1165)

Qy 4 GGTGCCACAGGGCAACACAGTAGTCCATCGACTACCGAAGACAGTTGACTGGGAGTG 63
Db 51 GlyAlaThrGlySerValTyPrTrpThrArgLeuLeuIleGlyHis 70
Qy 64 GGACGGGTAAACCCACTCGTTTCTGGTCATACCTGAGTGAGTCCGCCAGCCCCCTCTTAGGTAGA 123

Db 71 LysGlnValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGlyArg 90
:::|||||
QY 124 GACTATTGACCAAGTGGGAGCCAAATTTCTTTTGNACAGGGAACCAAGAGTGTCT 183
|||||
Db 91 AspLeuLeuThrLysLeuLysAlaGlnIleGlnPheSerAlaGluGlyProGlnValThr 110
|||||
QY 184 GCAATAACAACCTATCACTGTGTGACCTCCAAATTTAGATGACGAATATCGACTATAC 243
|||||
Db 111 TrpGlyGluArgProThrMetCysLeuValLeuAsnLeuGluGluTyArgLeuHis 130
|||||
QY 244 TCTCCCTAGTAAGCCTGTGATCAAAATATACAATTTCTGTTGGAACAGATTTCCTCCCAAGCC 303
|||||
Db 131 GluLysProVal---ProSerSerIleAspProSerTrpLeuLeuPheProThrVal 149
|||||
QY 304 TGGGAGAAACCGCGGATGGTTTGGCAAGCGAGTTCCCTCCCAAGATTATTCACACTG 363
|||||
Db 150 TrpAlaGluArgAlaGlyMetGlyLeuAlaAsnGlnValProProValValGluLeu 169
|||||
QY 364 AAGGCCAGTGCACACCACTGTCTAGTACAGACAGTACCTCTTGTAGTAAGAAGCTCAAGAA 423
|||||
Db 170 ArgSerGlyAlaSerProValAlaValArgGlnTyrProMetSerLysGluAlaArgGlu 189
|||||
QY 424 GGAATTCGGCGCGATGTCCTCAAGATTAATCCAACAGGGCATCTCTAGTTCTCTGTCCTCAATCT 483
|||||
Db 190 GlyIleArgProHisIleGlnLysPheLeuAspLeuGlyValLeuValProCysArgSer 209
|||||
QY 484 CCTCGGAATATCTCCCTGCTACCGGTTAGAAAGCTTGGAGCTATGACTATCGACAGTA 543
|||||
Db 210 ProTrpAsnThrProLeuLeuProValLysProGlyThrAsnAspTyrArgProVal 229
|||||
QY 544 CAGAGCTTGAGAGGTCAATAAACGGGTGAGGATATACCCCAACAGTCCCGAACCT 603
|||||
Db 230 GlnAspLeuArgGluIleAsnLysArgValGlnAspIleHisProThrValProAsnPro 249
|||||
QY 604 TATAACCTCTTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATTTGGACTTAAG 663
|||||
Db 250 TyrAsnLeuLeuSerSerLeuProProSerTyrThrTrpTyrSerValLeuAspLeuLys 269
|||||
QY 664 GATGCTTCTTCTGCTGAGATTACACCCCTAGCCACCAACCACTTTTGGCTTGAATGG 723
|||||
Db 270 AspAlaPhePheCysLeuArgLeuHisProAsnSerGlnProLeuPheAlaPheGluTrp 289
|||||
QY 724 AGAGATCCAGGTACCGGAGAACCGGCGAGCTCACCTGACCGCTGCTCCCAAGGGTTC 783
|||||
Db 290 LysAspProGluLysGlyAsnThrGlyGlnLeuThrTrpThrArgLeuProGlnGlyPhe 309
|||||
QY 784 AAGAACTCCCGCACCATCTTTGACGAAGCCCTTACACAGAGACCTGGCCCAACTTCAGGATC 843
|||||
Db 310 LysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaProPheArgAla 329
|||||
QY 844 CAACACCTCAGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCCACC 903
|||||
Db 330 LeuAsnProGlnValValLeuLeuGlnTyrValAspAspLeuLeuValAlaAlaProThr 349
|||||
QY 904 AAACAGGACTGTTAGAGGCACGAGGACCTACTGCTGGAATCTCTGACCTAGGCTAC 963
|||||
Db 350 TyrGluAspCysLysGlyThrGlnLysLeuLeuGlnLeuSerLysLeuGlyTyr 369
|||||
QY 964 AGAGCTCTGCTGAAGGCCACAGATTTGACGAGAGAGGTAACTACTTTGGGGTACAGT 1023
|||||
Db 370 ArgValSerAlaLysLysAlaGlnLeuCysGlnArgGluValThrTyrLeuGlyTyrLeu 389
|||||
QY 1024 TTGCGGAGCGGCAGCGATGGCTGACGAGGACGAGGACGAGAACTGTGTATCGATACCG 1083
|||||
Db 390 LeuLysGluGlyLysArgTrpLeuThrProAlaArgLysAlaThrValMetLysIlePro 409
|||||
QY 1084 GCCCAACACACGACCAACAAATGAGAGATTTTGGGACAGCTGGATTTTGACACTG 1143
|||||
Db 410 ValProThrThrProArgGlnValArgGluPheLeuGlyThrAlaGlyPheCysArgLeu 429
|||||
QY 1144 TGGATCCCGGGGTTTGGACCTTAGCAGGCCCACTCTACCCCTCAACCCCAAGAAAAAGG 1203
|||||

Db 430 TrpIleProGlyPheAlaSerLeuAlaAlaProLeuTyrProLeuThrLysGluSerIle 449
QY 1204 GAATTCCTCTGGGCTCTCTGAGCACACAGAGCATTTGATGCTATCAAAAAGCCCTGCTG 1263
|||||
Db 450 ProPheIleTrpThrGluGluHisGlnAlaPheAspHisIleLysLysAlaLeuLeu 469
|||||
QY 1264 AGGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCCCTTTACCTTTATGTGTGATGAG 1323
|||||
Db 470 SerAlaProAlaLeuAlaLeuProAspLeuThrLysProPheThrLeuTyrIleAspGlu 489
|||||
QY 1324 CGTAAGGAGTAGCCCGGAGTTTAAACCAACCCCTAGGACCATCGAAGACCTGTC 1383
|||||
Db 490 ArgAlaGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgProVal 509
|||||
QY 1384 GCCTACCTGTCAAAAGAGCTCGATCCTGTAGCCAGTGTGTGGCCCATATGCTCAAGGCT 1443
|||||
Db 510 AlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProThrCysLeuLysAla 529
|||||
QY 1444 ATCGCAGCTGTGGCCATCTGGTCAAGACGCTGACAAATTGACTTTGGGACAGAAATATA 1503
|||||
Db 530 ValAlaAlaValAlaLeuLeuLysAspAlaAspLysLeuThrLeuGlyGlnAsnVal 549
|||||
QY 1504 ACTGTAATAGCCCCCATGATTTGGAGAACTGCTTCGCGACGCCCCAGACCGATGATG 1563
|||||
Db 550 ThrValIleAlaSerHisSerLeuSerIleValArgGlnProProAspArgTrpMet 569
|||||
QY 1564 ACCAACCCCGCATGACCCACTATCAAGCTGCTCTCACAGAGAGGGTCACTTCGCT 1623
|||||
Db 570 ThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuAsnGluArgValSerPheAla 589
|||||
QY 1624 CCACGACCGCTCTCAACCCCTGCCACTTCTTGCTGGAAGACTGTATGAACCACTGACT 1683
|||||
Db 590 ProProAlaValLeuAsnProAlaThrLeuLeuProValGluSerGluAlaThrProVal 609
|||||
QY 1684 CATGATTCCTCAACTATTGATTTAGGAGACTGGGTCCGCAAGGACCTTACAGACATA 1743
|||||
Db 610 HisArgCysSerGluIleLeuAlaGluThrGlyThrArgArgAspLeuGluAspGln 629
|||||
QY 1744 CCGCTGACTGAGAAGTGTAACTGTTCACTCACGAGACAGTATGTGGTGAAGGT 1803
|||||
Db 630 ProLeuProGly---ValProThrTrpTyrThrAspGlySerSerPheIleThrGluGly 648
|||||
QY 1804 AAGAGATGGTGGGCGGCGTGTGACGGGACCCGACGATCTGGCCGACGAGCCTG 1863
|||||
Db 649 LysArgArgAlaGlyAlaProIleValAspGlyLysArgThrValTrpAlaSerSerLeu 668
|||||
QY 1864 CCGAAGAACTTCAGCACAAAAGCTGAGTCTATGCGCTCAGCAAGCTTTCGGCTG 1923
|||||
Db 669 ProGluGlyThrSerAlaGlnLysAlaGluLeuValAlaLeuThrGlnAlaLeuArgLeu 688
|||||
QY 1924 GCCGAAGGGAATCCATAACATTTATACGACAGCAGTATGCTTTGCCACTGCACAC 1983
|||||
Db 689 AlaGluGlyLysAsnIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAlaHis 708
|||||
QY 1984 GTACATGGGCGCATCTATAAACAAGGGGTGTCTACCTCAGCAGGAGGGAATAAAG 2043
|||||
Db 709 IleHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyLysAspIleLys 728
|||||
QY 2044 AACAAAGAGGAAATTTCTAAGCCTATTAGAACCGTACATTTACCAAAAGGCTAGCTAAT 2103
|||||
Db 729 AsnLysGluGluIleLeuAlaLeuLeuGluAlaIleHisLeuProArgValAlaIle 748
|||||
QY 2104 ATACACTGTCTGGACATCACAAAGCTCAATATCTCATATCCAGAGAAACACAGATGCT 2163
|||||
Db 749 IleHisCysProGlyHisIleArgGlySerAsnProValAlaThrGlyAsnArgAla 768
|||||
QY 2164 CACCGGTTGGCAAGCAGGCGAGGCTGTAACTTCTGCTCTATATATAGAAATGCC 2223
|||||
Db 769 AspGluAlaAlaLysGlnAlaLeuSerThrArgValLeuAlaGlyThrThrLysPro 788
|||||
QY 2224 AAAGCCCCA---GAACCCAGACGACGTACACCTTAGAAGACTGGCAGAGATAAAG 2280
|||||
Db 789 GlnGluProIleGluProAlaGlnGluLysThrArgPro----- 801
|||||

QY	2281	ATAGACCAGTTCTCTGAGACTCCCGAAGGNCCTGCTATACCTCAGATGGGAAGAAATC	2340
DB	802	-----ArgGluLeu	804
QY	2341	CTGCCCCACAAGAAGGGTTAGAAATATGTCACACAGATACATCGTCTAAACCACCTAGGA	2400
DB	805	ThrProAspArg--GlyLeuGluPheIeIysArgLeuHisGlnLeuThrHisLeuGly	823
QY	2401	ACTAAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTTGAGGCTACCAAGGA	2460
DB	824	ProGluIysLeuLeuGlnLeuValAsnArgThrSerLeuLeuLeuIleProAsnLeuGlnSer	843
QY	2461	GTGGCTGACTCGGGTGGTCAAAACATTGTCCTGCCAGCTGGTGAATGCTTAATCCTCTTC	2520
DB	844	AlaValArgGluValThrSerGlnCysGlnAlaCysAlaMetThrAsnAlaValThrThr	863
QY	2521	AGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCAGCGCTCACTGGGAAGTGGAC	2580
DB	864	TyrArgGluThrGlyIysArgGlnArgIysAspArgProGlyValTyrTrpGluValAsp	883
QY	2581	TTCACTCAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGGTTTGTGTAGACACC	2640
DB	884	PheThrGluIleIysProGlyArgTyrGlyAsnIysTyrLeuLeuValPheIleAspThr	903
QY	2641	TTTTTCAGATGGGTAGAGGCTTATCTCTACTAAAGAAAGAGACTTCAACCGTGGTGGCTAAA	2700
DB	904	PheSerGlyTrpValGluAlaPheProThrLysThrGluThrAlaLeuIleValCysLys	923
QY	2701	AAAAATCTCGAAGAAAATTTTTTCCAAGATTTCGAATACCTAAGGTAAATAGGCTCAGACAAT	2760
DB	924	LysIleLeuGluLeuLeuProArgPheGlyIleProLysValLeuGlySerAspAsn	943
QY	2761	GGTCCAGCTTTTGTGGCCCAAGTAAGTCAGGAGCTGGCCCAAGATATATGGGGATGTATGG	2820
DB	944	GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaThrGlnLeuGlyIleAsnTrp	963
QY	2821	AAACTGCATTGTGCATACAGACCCCAAGCTCAGGACAGCTAGAGAGGATGAATAGAAC	2880
DB	964	LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr	983
QY	2881	ATTAAAGAGACCTTACTAAATTGACCGCGAGACTGGCGCTTAATGATTGGATAGCTCTC	2940
DB	984	IleLysGluThrLeuThrLysLeuAlaLeuGluThrGlyGlyLysAspTrpValThrLeu	1003
QY	2941	CTGCGCTTTGTGCTTTTAGGGTTAGGAACACCCCTGGACAGTTGGCTGACCCCTTAT	3000
DB	1004	LeuProLeuAlaLeuLeuArgAlaArgAsnThrProGlyArgPheGlyLeuThrProTyr	1023
QY	3001	GAATTACTCTACGGGGACCCCGCCCAATTCGTAGAAATTCCTCTGTACATAGTCTGCAC	3060
DB	1024	GluIleLeuTyrGlyGlyProProIleLeuGluSerGlyGluThrLeuGlyProAsp	1043
QY	3061	GTGCTGCTTCCAGCGCTTTGTCTCTAGGCTCAAGGCACCTTCAGTGGGTGAGACAACGA	3120
DB	1044	AspArgPheLeuProValLeuPheThrHisLeuLysAlaLeuGluIleValArgThrGln	1063
QY	3121	CGCTGGAGGCACTCCGGAGGCGCTACTCAGGAGGAGGACACTTCGAGATCCCAATCGT	3180
DB	1064	IleTrpAspGlnIleLysGluValTyrLys--ProGlyThrValThrIleProHisPro	1082
QY	3181	TTCCAAGTGGGAGATTCAGTCTACGTTAGACGCCACCGTCAGGAAACCTTCGAGACTCG	3240
DB	1083	PheGlnValGlyAspGlnValLeuValArgArgHisArgProSerSerLeuGluProArg	1102
QY	3241	TGGAAGGGCCCTTATCTCGTACTTTTGACCAACCAACCGGCTGTGAAGTCGAAGGAATC	3300
DB	1103	TriLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValAspGlyIle	1122
QY	3301	TCCACCTGGATCATGCATCCACAGTTTAAACCGCGCCACCT-----CCCGATTCCGGG	3354
DB	1123	AlaAlaTrpValHisAlaSerHisLeuLysProAlaProProSerAlaProAspGluSer	1142

QY 3355 TGGAAAGCGGAAAAACACTGAATAATCCCTTAAGCTTCGCCTCATCGC 3402
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
1143 TrpGluLeuGlyThrAspHisProLeuLysLeuArgGlnValGarg 1158

RESULT 2
GNMVGW

HIV-1 retropepsin (EC 3.4.23.16) - AKV murine leukemia virus
N:Contains: nuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DN
C:Species: AKV murine leukemia virus
A:Note: host Mus spp. (mouse)
C>Date: 27-Nov-1985 #sequence revision 27-Nov-1985 #text_change 31-Dec-2004
C/Accession: B92995; B93448; A03957
R:Herr, W.

J. Virol. 49, 471-478, 1984
A>Title: Nucleotide sequence of AKV murine leukemia virus.
A/Reference number: A92995; MUID:84115072; PMID:6319746
A/Accession: B92995
A:Molecule type: genomic RNA
A:Residues: 1-1196 <HE>
A/Cross-references: UNIPROT:Q9EYM1; UNIPROT:Q83362; UNIPROT:Q9JBE2; UNIPROT:Q9JBE2; UNIPROT:Q9JBE2; UNIPROT:Q9JBE2
F:3-102/Product: retropepsin #status predicted <RP>
P:27/Active site: Asp (shared with dimeric partner) #status predicted

C/Comment: This protein is synthesized as a gag-pol polyprotein.
C/Comment: The pol polyprotein contains reverse transcriptase and possibly a n
C/Genetics:
A:Gene: pol
C/Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polyprote
F:3-102/Product: retropepsin #status predicted <RP>
P:27/Active site: Asp (shared with dimeric partner) #status predicted

Alignment Scores:
Pred. No.: 1.24e-248 Length: 1196
Score: 3836.00 Matches: 722
Percent Similarity: 77.4% Conservative: 166
Best Local Similarity: 62.8% Mismatches: 246
Query Match: 62.0% Indels: 14
DB: 1 Gaps: 9

US-10-723-552-3_COPY_2307_5741 (1-3435) x GNMVGW (1-1196)

QY 4 GGTCGCCACGGCAACAACATATCCATCGACTCCCGAAGAACAAGTGTGAGTG 63
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
51 GlyAlaThrGlyGlyLysArgTrpThrAspArgLysValHisLeuAlaThr 70

QY 64 GGACGGGTAAACCCTGTTTCTGTGTATCTAGTCGCCAGCACCCCTCTTAGGTAGA 123
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
71 GlyLysValThrHisSerPheLeuHisValProAspCysProTyProleuLeuGlyArg 90

QY 124 GACTATTATGACCAAGATGGGAGCACAAATTTCTTTTGAA---CAAGGGAAACCAAGAGTG 180
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
91 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValVal 110

QY 181 TCTGCAATAAACAAACCTATCACCTGTGTGTGACCCCTCCAAATAGATGACGAATACGACTA 240
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
111 GlyProLysGlyGlnProLeuGlnValLeuThrLeuAsnLeuGluAspGluTrgArgLeu 130

QY 241 TACTCTCCCCTAGTAAAGCCCTGATCAAATATA---CAATTCTGGTGTGAACAGTTTCCC 297
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
131 TyrGluThrSerAlaGluProGluValSerProGlySerThrTrpLeuSerAspPhePro 150

QY 298 CAAGCCTGGGAGAAAACCGCAGGATGGTGTGGCAAGACAGTTCCTCCCAACAAGTTATT 357
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
151 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProleuIleIle 170

QY 358 CAACTGAAGGCCAGTCGCACACCACTGTCAGTCTAGACAGTACCTCCCTTGGTAAAGAGCT 417
DB ||||| :||||:||||:||||:||||:||||:||||:||||:||||:
171 ProLeuLysAlaThrSerThrProValSerIleLysGlnTrpMetSerGlnGluAla 190

Db 909 ProGlySerHisTrpGluIleAspPheThrGluValLysProGlyLeuTyrGlyTyrLys 928
Qy 2617 TACCTATTGGTTTGTAGACACCTTTTTCAGGATGGGTAGAGGCTTATCTTACTAAGAAA 2676
Db 929 TyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaPheProThrLysArg 948
Qy 2677 GAGACTTCAACGGTGGTGGCTTAAATAATATCTGGAAGAAATTTTCCAAAGATTGGGAATA 2736
Db 949 GluThrAlaArgValLysLysLysLeuGluIlePheProArgPheGlyMet 968
Qy 2737 CCTAAGGTAAATAGGTCCAGCAATGGTCCAGCTTTGTGGCCAGGTAAAGTCAGGAGCTG 2796
Db 969 ProGlnValLeuGlySerAspAsnGlyProAlaPheThrSerGlnValSerGlnSerVal 988
Qy 2797 GCCAAGATATTGGGAGTATGATTGAAACTGCATTTGTGCATACACAGCCCAAGCTCAGGA 2856
Db 989 AlaAspLeuLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSerGly 1008
Qy 2857 CAGGTAGAGAGGATGAATAGAACCATTAAGAGACCCCTTACTAAATGACCGCGAGACT 2916
Db 1009 GlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrLeuAla 1028
Qy 2917 GGCCTTAATGATTGATAGCTCTCTGCTCCCTTTGTGCTTTTATAGGTTAGCAACCCCT 2976
Db 1029 GlyThrArgAspTrpValLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThrPro 1048
Qy 2977 GGACAGTTTGGGCTGACCCCTATGAAATCTCTACGGGGACCCGCCCTTGGTGA 3036
Db 1049 GlyProHisGlyLeuThrProTyrGluIleLeuTyrGlyAlaProProLeuValAsn 1068
Qy 3037 ATTGCTTCTGACATGCTGCTGCTGCTTTCCAGCGCTTTGTTCTTCTAGGCTCAAG 3096
Db 1069 PheHisAspProAspMetSerGluLeuThrAsnSerProSerLeuGlnAlaHisLeuGln 1088
Qy 3097 GCACCTGAGTGGGTGAGACACGAGCGTGGAGCACTCCGGGAGGCTTACTCAGGAGGA 3156
Db 1089 AlaLeuGlnThrValGlnArgGluIleTrpLysProLeuAlaGluAlaTyrArgAspGln 1108
Qy 3157 GGAGCTTGCAG---ATCCACATCGTTTCCAGTGGGAGATTTCAGTCTAGTTAGACGC 3213
Db 1109 LeuAspGlnProValIleProHisProPheArgIleGlyAspSerValTrpValArgArg 1128
Qy 3214 CACCGTCAGGAAACCTCGAGACTCGGTGGAGGGCCCTTATCTCGTACTTTTGACACACA 3273
Db 1129 HisGlnThrLysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThrThr 1148
Qy 3274 CCAACGCTGTGAAGTCGAAGGAATCTCCACCTGGATCCATGCTCCAGTCCAGTTAAACCG 3333
Db 1149 ProThrAlaLeuLysValAspGlyIleSerAlaTrpIleHisAlaAlaHisValLysAla 1168
Qy 3334 GCG-----CCACCTCCCGATTCCGGGTGGAAACCGCAAGACTGAAATCCCTTAAG 3387
Db 1169 AlaThrThrProProIleLysProSerTrpArgValGlnArgSerGlnAsnProLeuLys 1188
Qy 3388 CTTCCCTCCATCCGCGTGGTTCCT 3411
Db 1189 IleArgLeuThrArgGlyAlaPro 1196

RESULT 3

S35475

pol polyprotein - Friend murine leukemia virus (strain PVC-211)

N;Contains: nuclease; proteinase; reverse transcriptase

C;Species: Friend murine leukemia virus

A;Variety: strain PVC-211

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C;Accession: S35475

R;Remington, M.P.; Hoffman, P.M.; Ruscelli, S.K.; Masuda, M.

Nucleic Acids Res. 20, 3249, 1992

A;Title: Complete nucleotide sequence of a neuropathogenic variant of Friend murine leuk

A;Reference number: S35474; MUID:92319660; PMID:1620621

A;Accession: S35475

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: genomic RNA

A;Residues: 1-1204 <REM>

A;Cross-references: UNIPROT:P26808; UNIPARC:UPI0000131F06; EMBL:M93134; NID:G331898; P1

A;Experimental source: strain PVC-211

A;Note: the authors translated the stop codon at position 5 as Gln

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992

C;Genetics:

A;Gene: pol

C;Superfamily: pol polyprotein

C;Keywords: polyprotein; reverse transcriptase

Alignment Scores:

Pred. No.:	1.85e-247	Length:	1204
Score:	3818.50	Matches:	723
Percent Similarity:	76.2%	Conservative:	157
Best Local Similarity:	62.6%	Mismatches:	244
Query Match:	61.8%	Indels:	31
DB:	2	Gaps:	11

US-10-723-552-3_COPY_2307_5741 (1-3435) x S35475 (1-1204)

Qy 4 GGTGCCACAGGGCAACAACAGTATCCATGGACTACCCGAGAACAGTTGACTTGGAGTG 63
Db 56 GlyAlaThrGlyLysArgTyrArgTrpThrThrAspArgArgValHisLeuAlaThr 75
Qy 64 GGCGGTAAACCCACTCGTTTCTGTCATACCTGAGTCCCGCAGCACCCCTCTTAGGTAGA 123
Db 76 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 95
Qy 124 GACTTATGACCAAGATGGGAGCACAAATTTCTTTGAA---CAAGGGAACCAAGAGTG 180
Db 96 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValVal 115
Qy 181 TCTGCAATAACAAACCTATCCTGCTGCTGCACCTCCCAATTAGATAGCAATATCGACTA 240
Db 116 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluTyrArgLeu 135
Qy 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATA---CAATTCTGTTGGAAACAGTTTCCC 297
Db 136 HisGluThrSerLysGlyProAspValProLeuGlySerThrTrpLeuSerAspPhePro 155
Qy 298 CAAGCTGGCGAGAACCCGAGGATGGTGTGGCAAGCAAGTTCCCCCAAGATTATT 357
Db 156 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 175
Qy 358 CAAGTGAAGCCAGTGCACACAGTCTAGTCAGACAGTACCCTTGAGCTAAAGAGCT 417
Db 176 ProLeuArgAlaAlaSerThrProValSerIleLysGlnTyrProMetSerArgGluAla 195
Qy 418 CAAGAAGGAATTCGGCCGATGTCCAAAGATTAAATCAACAGGGCATCCTAGTTCTGTC 477
Db 196 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 215
Qy 478 CAATCTCCCTGGAAATCTCCCTGCTACCGGTAGAAAGCCTGGAGCTAATAGTATCGA 537
Db 216 GlnSerProTrpAsnThrProLeuLeuProValLysProGlyThrAsnAspTyrArg 235
Qy 538 CCAGTACAGACTTGAGAGAGGTCAATAACGGGTGCAGGATATACACCAACAGTCCCG 597
Db 236 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 255
Qy 598 AACCTTTATAACCTTGTGTGCTCTCCCAACCCCAACCGAGCTGGTATACAGATTATGAC 657
Db 256 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp 275
Qy 658 TTAAGAGATGCTTCTTCTGCTCGCTGAGATTACACCCCACTAGCCCAACCACTTTTGCCTTC 717
Db 276 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnSerLeuPheAlaPhe 295
Qy 718 GAATGAGAGATCCAGGTACGGGAAGAACCGGACGCTCACCTGGACCCCACTGCCCA 777
Db 296 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 315

Db 1027 ThrLysLeuThrLeuAlaThrGlySerArgAspTrpValLeuLeuLeuProLeuAlaLeu 1046
 Qy 2956 TTTAGGGTTAGGAACACCCCTGGACAGTTGGGCTGACCCCTATGAATTACTCTACCGG 3015
 Db 1047 TyrArgAlaAraGlnThrProGlyProHisGlyLeuThrProTyrGluLeuTyrGly 1066
 Qy 3016 GGACCCGCCCATTTGGTGAATAATGCTCTGTGTACATAGTGTGACGCTGCTTCCCG 3075
 Db 1067 AlaProProProLeuValAsnPheProAspProAspMetAlaLysValThrHisAsnPro 1086
 Qy 3076 CTTTGTCTCTAGCTCAAGCACTTGGTGGGTGACACCAAGCGGTGAGGCACTC 3135
 Db 1087 SerLeuGlnAlaHisLeuGlnAlaLeuTyrLeuValGlnHisGluValTrpArgProLeu 1106
 Qy 3136 CGGAGGCGCTACTCAGGAGGAGGACTTGCAG---ATCCACATCGTTTCCAAAGTGGGA 3192
 Db 1107 AlaAlaAlaTyrGlnGlnLeuAspArgProValValProHisProPheArgValGly 1126
 Qy 3193 GATTGAGTCTAGTTAGACGCCAGCGTGCAGAACTCGAGACTCGGTGGAAGGCGCT 3252
 Db 1127 AspThrValTrpValArgHisGlnThrLysAsnLeuGluProArgTrpLysGlyPro 1146
 Qy 3253 TATCTCGPACTTTGACACCAACCGCTGTGAAGTTCGAAGAACTCCACCTGGATC 3312
 Db 1147 TyrThrValLeuLeuThrThrProThrAlaLeuLysValAspGlyIleAlaAlaTrpIle 1166
 Qy 3313 CATGCATCCACGTTAAACCGCGC-----CCACCTCCCGATTCCGGGTGG 3357
 Db 1167 HisAlaAlaHisValLysAlaAlaAspThrLysIleGluProProSerGluSerThrTrp 1186
 Qy 3358 AAAGCCGAAAGACTGAATAATCCCTTAAGCTTCGCTCCCATCGC 3402
 Db 1187 ArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArg 1201

RESULT 4

pol polyprotein - Friend murine leukemia virus (strain FB29)
 N;Contains: nuclease; proteinase; reverse transcriptase
 C;Species: Friend murine leukemia virus
 A;Variety: strain FB29
 C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
 C;Accession: S70393
 R;Perrymann, S.; Nishio, J.; Chesebro, B.
 Nucleic Acids Res. 19, 6950, 1991
 A;Title: Complete nucleotide sequence of Friend murine leukemia virus, strain FB29.
 A;Reference number: S70393; MUID:92107687; PMID:1762923
 A;Accession: S70393
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: genomic RNA
 A;Residues: 1-1204 <PER>
 A;Cross-references: UNIPROT:P26809; UNIPARC:UPI0000131P05; EMBL:Z11128; NID:G61547; PIDN

A;Experimental source: strain FB29
 A;Note: the authors translated the stop codon at position 5 as Gln
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
 C;Genetics:
 A;Gene: pol
 C;Superfamily: pol polyprotein
 C;Keywords: polyprotein; reverse transcriptase

Alignment Scores:

Pred. No.: 7,46-247 Length: 1204
 Score: 3809.50 Matches: 720
 Percent Similarity: 76.6% Conservative: 160
 Best Local Similarity: 62.7% Mismatches: 250
 Query Match: 61.6% Indels: 19
 DB: 2 Gaps: 10

US-10-723-552-3_COPY_2307_5741 (1-3435) x S70393 (1-1204)

Qy 4 GGTGCCACGGGCACACAGTATCCATGACTACCCGAGACAGCTGACTGGGAGTG 63
 Db 56 GlyAlaThrGlyGlyLysArgTyrArgTrpThrAspArgValHisLeuAlaThr 75

Qy 64 GGACGGGTAAACCCACTCGTTCTTCTGTCATACCTGAGTGCCCGACGACCCCTCTTAGGTAGA 123
 Db 76 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 95
 Qy 124 GACTTATTTGACCAAGATGGGAGGACAAATTTCTTTTGAA----CAAGGGAAACCGAAGTG 180
 Db 96 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValVal 115
 Qy 181 TCTCAAAATAACAACCTATCCTCCTGTGTGACCTCCAAATTAGATGACGAATATCGACTA 240
 Db 116 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluTyrArgLeu 135
 Qy 241 TACTCTCCCTAGTAAAGCCCTGATCAAAATATA---CAATTTCTGTTGGAAACAGTTTCCC 297
 Db 136 HisGluThrSerLysGlyProAspValProLeuGlySerThrTrpLeuSerAspPhePro 155
 Qy 298 CAAGCCTGGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCTCCCAAGATTATT 357
 Db 156 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 175
 Qy 358 CAAGCTGAAGCGCAGTGCACACAGTGTCTAGTCAGACAGTACCCCTTGAGTAAAGAGCT 417
 Db 176 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrProMetSerGlnGluAla 195
 Qy 418 CAAGAAAGGAATTCGGCCGCATGTCCAAAGATTAAATCCAAACAGGCGCATCTAGTTCTCTGC 477
 Db 196 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 215
 Qy 478 CAATCTCCCTGGGAATATCTCCCTGCTACCGGTTAGAAAGCTTGGGACTAATGACTATCGA 537
 Db 216 GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 235
 Qy 538 CCAAGTACAGACTTTCAGAGAGGTCATAAACCAGGTGCAGGATATACACCCCAACAGTCCCG 597
 Db 236 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 255
 Qy 598 AACCTTTATAACCTCTTGTGTGTCTCTCCACCCCAACGAGAGCTGGTATACAGTATTGGAC 657
 Db 256 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTyrTrpThrValLeuAsp 275
 Qy 658 TTAAGAGATCCCTTCTCTGCTGAGATTACACCCCACTAGCCCAACACACTTTTTCCTTC 717
 Db 276 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnSerLeuPheAlaPhe 295
 Qy 718 GAATGGAGAGATCCAGGTACGCGAAGAACCGGGCAGCTCACCTGGACCCCGACTCCCCCAA 777
 Db 296 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 315
 Qy 778 GGGTTCAAGAACTCCCGACCATCTTTTGAAGAGCCCTACACAGAGACCTGGCCCACTTC 837
 Db 316 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 335
 Qy 838 AGATCCAAACACCTCAGGTGACCTCTCCAGTACGTGATGACCTGCTTCTGGCGGGA 897
 Db 336 ArgIleGlnHisProAspLeuIleLeuLeuGlnTyrValAspAspLeuLeuLeuAlaAla 355
 Qy 898 GCCACCAACACAGACTCTTAGAAGGACCAAGGACCTACTGCTGGAATTTGCTGACCTA 957
 Db 356 ThrSerGlnLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAspLeu 375
 Qy 958 GGTACAGAGCTCTCTGAAGAGGCGCCAGATTGTCAGAGAGAGAGGTAAACATACTCTGGG 1017
 Db 376 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 395
 Qy 1018 TACAGTTTGGGACCGGGCAGCGATGGCTGACCGAGGACGGAAGAAACCTGTAGTCCAG 1077
 Db 396 TyrLeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly 415
 Qy 1078 ATACCGGCCCAACACAGCCAAACAAATGAGAGAGTTTGGGACAGCTGGATTTGCT 1137
 Db 416 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 435
 Qy 1138 AGACTGTGGATCCCGGGTTTTCGACCTTAGCAGCCCACTCTACCGCGCTAACCAAGAA 1197

Db 436 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 455
Qy 1198 AAAGGGGAATTCCTCTGGGCTCTGAGCACAGAGGCAATTTGATGCTATCAAAAAGGCC 1257
Db 456 GlyThrLeuPheGluTrpGlyProAspGlnGlnLysAlaLysAlaLysGlnAla 475
Qy 1258 CTGCTGAGCGCACCTGCTCGCCCTCCCTGACGTTAACTTAACCTTTACCTTTATGTCG 1317
Db 476 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 495
Qy 1318 CATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCCCTAGGACCACTGAGAGA 1377
Db 496 AspGlnLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTyrPheArg 515
Qy 1378 CTTGTGCTGCTACCTGTCAAAGAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGCTG 1437
Db 516 ProValAlaLysLeuSerLysLysLeuAspProValAlaAlaGlyTyrProProCysLeu 535
Qy 1438 AAGGCTATCGAGCTGTGGCCATCTGTCAGGACGCTGACAAATTCACCTTTGGGACG 1497
Db 536 ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln 555
Qy 1498 AATATACTGTAATAGCCCCCATCGATTGGAGAACATCGTTGCGCAGCCGCCAGACCGA 1557
Db 556 ProLeuValIleLeuAlaProHisAlaValGluAlaLeuValLysGlnProProAspArg 575
Qy 1558 TGGATGACCAACGCCCGCATGACCCATATCAAGCCTGCTTC--ACAGAGAGGTC 1614
Db 576 TrpLeuSerAsnAlaArgMetThrHisTyrGlnAlaLeuLeuLeuAspThrAspArgVal 595
Qy 1615 ACGTTCGCTCCACAGCCGCTGCAACCTCGCCTCTTCCTGCTGAGAGACTGATGAA 1674
Db 596 GlnPheGlyProIleValAlaLeuAsnProAlaThrLeuLeuPro---LeuProGluGlu 614
Qy 1675 CCAGTCACTCATGATTGTCATCACTATTTGATTGAGGAGACTGGGGTCCGCAAGACCTT 1734
Db 615 GlyLeuGlnHisAspCysLeuAspIleLeuAlaGluAlaHisGlyThrArgProAspLeu 634
Qy 1735 ACAGACATACCGCTGACTGGAGAAGTGTAACTGTTGCTACGACGAGAGCTATGTG 1794
Db 635 ThrAspGlnProLeuProAspAlaAspHisThrTrpTyrThrAspGlySerSerPheLeu 654
Qy 1795 GTGAGAGTAAGAGATCGCTGGCGGGGGTGGTGGACGGGACCGCACGACTGGGCC 1854
Db 655 GlnGluGlyGlnArgLysAlaGlyAlaAlaValThrThrGluThrGluValValTrpAla 674
Qy 1855 AGCAGCTGCGGGAAGGAACCTTCAGCACAAAGGCTGAGCTCATGGCCCTCACGCAAGCT 1914
Db 675 LysAlaLeuProAlaGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAla 694
Qy 1915 TTGGCGTGGCGGAGGGAATTCATAAATTTATACGACAGAGATATGCTTTGCG 1974
Db 695 LeuLysMetAlaGluGlyLysLysLeuAsnValTyrThrAspSerArgTyrAlaPheAla 714
Qy 1975 ACTGCACAGCTACATGGGCCATCTATAACAAAGGGGTTGCTTACCTCAGCAGGAGG 2034
Db 715 ThrAlaHisIleHisGlyGluIleTyrArgArgGlyLeuLeuThrSerGluGlyLys 734
Qy 2035 GAAATAAAGAACAAAGGAATTTAAAGCCTATTAGAAGCGGTACATTTTACCAAAAAGG 2094
Db 735 GluIleLysAsnLysAspGluIleLeuAlaLeuLeuLysAlaLeuPheLeuProLysArg 754
Qy 2095 CTAGCTATTATACATGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAC 2154
Db 755 LeuSerIleIleHisCysProGlyHisGlnLysGlyAsnArgAlaGluAlaArgGlyAsn 774
Qy 2155 CAGATGGCTGACCGGTTGCCAAGCAGCAGC-----CAGGTTGTTAAC 2199
Db 775 ArgMetAlaAspGlnAlaAlaArgGluValAlaThrArgGluThrProGluThrSerThr 794
Qy 2200 CTTCTGCTATAATAGAAATGCCCAAAGCCCCAGAACCCAGACGACGATACACCTAGAA 2259
Db 1153 ThrProThrAlaLeuLysValAspGlyIleAlaAlaTrpIleHisAlaAlaHisValLys 1172

Db 795 LeuLeuIleGluAsnSerAlaProTyrThrHisGlu---HisPheHisTyrThrValThr 813
Qy 2260 GACTGGCAAGAGATAAAAAAGATAGACAGTTCTCTGAGACTCCGGAGGACCTGCTAT 2319
Db 814 AspileLysAspLeuThrLysLeuGlyAlaThrTyrAspAspAlaLys---LysCysTrp 832
Qy 2320 ACCTCAGATGGAGGAAATCTCGCCCAACAAAGAGGTTAGAAATATGTCACACAGATA 2379
Db 833 ValTyrGlnGlyLysProValMetProAspGlnPheThrPheGluLeuLeuAspPheLeu 852
Qy 2380 CATCGTTAAACCCACCTAGGAACATAAACACCTGCAGCAGTTGGTCAGAACATCC----- 2433
Db 853 HisGlnLeuThrHisLeuSerPheSerLysThrLysAlaLeuLeuGluAlaArgAsnTyrCys 872
Qy 2434 CTTTATCATGTTCTGAGGCTACACAGGAGTGGCTGACTCGGTGGTCAACATGTTGTGCC 2493
Db 873 ProTyrTyrMetLeuAsnArgAspArgThrLeuLysAspileThrGluThrCysGlnAla 892
Qy 2494 TGCAGCTGGTTAATGCTAATCTTCAGAAATGCTCCAGGAGAGACTAAAGGGGAGC 2553
Db 893 CysAlaGlnValAlaAlaSerLysSerAlaValLysGlnGlyThrArgValArgGlyHis 912
Qy 2554 CACCAGCGCTCACTGGGAAGTGCATCTCACTCAGGTAAAGCCGGCTAAATACGGAAAC 2613
Db 913 ArgProGlyThrHisTrpGluIleAspPheThrGluValLysProGlyLeuTyrGlyTyr 932
Qy 2614 AAATACCTATTTGTTTGTAGACACCTTTTCAGGATGGGTAGAGCTTATCCTACTAAG 2673
Db 933 LysTyrLeuLeuValPheIleAspThrPheSerGlyTrpValGluAlaPheProThrLys 952
Qy 2674 AAAGAGACTTCAACCGTGGTGGCTTAAAAAAATCTCGAAGAAATTTTCCAAAGATTGGA 2733
Db 953 LysGluThrAlaLysValValThrLysLysLeuLeuGluIlePheProArgPheGly 972
Qy 2734 ATACCTAAGTAAATAGGTCAGACAACTGGTCCAGCTTTTGTGTCGCCAGGTAAAGTCAGGA 2793
Db 973 MetProGlnValLeuGlyThrAspAsnGlyProAlaPheValSerLysValSerGlnThr 992
Qy 2794 CTGGCCAGATATTGGGATTTGATTGGAATCTGCAATCTGTCATACAGACCCCAAGCTCA 2853
Db 993 ValAlaAspLeuLeuGlyValAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer 1012
Qy 2854 GGACAGCTAGAGAGATAGAAATAGAACCATTAAGACAGACCTTACTATAATTCACCGCGAG 2913
Db 1013 GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrLeuAla 1032
Qy 2914 ACTGCGCTTAATGATAGATAGCTCTCTGCGCTTTGTGCTTTTGTAGGTTAGGAACACC 2973
Db 1033 ThrGlySerArgAspTrpValLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThr 1052
Qy 2974 CTTGCACAGTTTGGCTGACCCCTATATGATTTACTCTACGGGGAGCCCCCTATTGGTA 3033
Db 1053 ProGlyProHisGlyLeuThrProTyrGluIleLeuTyrGlyAlaProProLeuVal 1072
Qy 3034 GAAATGCTCTGTACATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3093
Db 1073 AsnPheProAspProAspMetAlaLysValThrHisAsnProSerLeuLeuAlaHisLeu 1092
Qy 3094 AAGCAGCTTGTAGTGGTGAGAACAGAGCGGTGGAGGCAACTCCGGGAGGCCCTACTCAGGA 3153
Db 1093 GlnAlaLeuTyrLeuValGlnHisGluValTrpArgProLeuAlaAlaAlaTyrGlnGlu 1112
Qy 3154 GGAGGAGCTTGCAG---ATCCACATCGTTTCCAAAGTGGAGATTCAGTCTAGCTTGA 3210
Db 1113 GlnLeuAspArgProValValProHisProPheArgValGlyAspThrValTrpValArg 1132
Qy 3211 CGCCACCGTGCAGAAACCTCGAGACTCGGTGGAGGGCCCTTATCTCGTACTTTTGACC 3270
Db 1133 ArgHisGlnThrLysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThr 1152
Qy 3271 ACACAAACGGCTGTGAAGTCAAGGAATCTCCACCTGGATCCATGTCATCCACGTTAAA 3330
Db 1153 ThrProThrAlaLeuLysValAspGlyIleAlaAlaTrpIleHisAlaAlaHisValLys 1172

Qy 3331 CCGCG-----CCACCTCCGATTGCGGTGGAAGACCGAAGAACTGAA 3375
 Db 1173 AlaAlaaspThrArglleGluProSerGluSerThrTrpArgValGlnArgSerGln 1192
 Qy 3376 AATCCCTTAAGCTTCGCTCCATCGC 3402
 Db 1193 AsnProLeuIysIleArgLeuThrArg 1201

RESULT 5
 GNMVM7
 HIV-1 retropepsin (EC 3.4.23.16) - baboon endogenous virus (strain M7)
 N;Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
 C;Species: baboon endogenous virus
 A;Note: host Papio sp. (baboon)
 C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C;Accession: J0261
 R;Kato, S.; Matsuo, K.; Nishimura, N.; Takahashi, N.; Takano, T.
 Jpn. J. Genet. 62, 127-137, 1987
 A;Title: The entire nucleotide sequence of baboon endogenous virus DNA: a chimeric genom
 A;Reference number: J0260
 A;Molecule type: DNA
 A;Residues: 1-1189 <KAT>
 A;Cross-references: UNIPROT:P10272; UNIPARC:UPI0000131EC7; GB:X05470
 C;Comment: This protein is synthesized as a gag-pol polyprotein.
 C;Genetics:
 A;Gene: pol
 C;Superfamily: pol polyprotein
 C;Keywords: aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase; revers
 F;1-120/Product: retropepsin #status predicted <RT>
 F;3-102/Product: retropepsin #status predicted <RT>
 F;121-797/Product: RNA-directed DNA polymerase #status predicted <REV>
 F;798-1189/Product: endonuclease #status predicted <EDB>
 F;27/Active site: Asp (shared with dimeric partner) #status predicted

Alignment Scores:
 Pred. No.: 2,17e-246 Length: 1189
 Score: 3802.50 Matches: 730
 Percent Similarity: 76.4% Conservative: 148
 Best Local Similarity: 63.5% Mismatches: 244
 Query Match: 61.5% Indels: 27
 DB: 1 Gaps: 12

US-10-723-552-3_COPY_2307_5741 (1-3435) x GNMVM7 (1-1189)

Qy 4 GGTGCCACGGCAACACAGTAGTATCCATGCTACCCGAGAACAGTTGACTGGGAGTG 63
 Db 51 GlyAlaThrGlyArgIysMetHisLysIleTrpThrAsnArgArgThrValAsnLeuGlyGln 70
 Qy 64 GGACGGGTAAACCCACTCGTTTCTGGTCATACCTGAGTGGCCAGCACCCCTCTTAGGTAGA 123
 Db 71 GlyMetValThrHisSerPheLeuValValProGluCysProTyrProLeuLeuGlyArg 90
 Qy 124 GACTTATGACCAAGATGGAGACAAATTCCTTTT---GAACAAGGAAACAGAAAGTG 180
 Db 91 AsnLeuLeuThrIysLeuGlyAlaGlnIleHisPheSerGluAlaGlyAlaGlnValLeu 110
 Qy 181 TCTGCAATAACAACTATCAGTGTGCTGCTGCTCCAAATAGATGACGAATATCGACTA 240
 Db 111 AspArgaspGlyGlnProIleGlnIleLeuThrValSerLeuGlnAspGluHisArgLeu 130
 Qy 241 TACTCT---CCCCTAGTAAAG-----CCTGATCAAAATATACAAATTCCTGGTTGGAAACAG 291
 Db 131 PheAspIleProValThrThrSerLeuProAsp-----ValTrpLeuGlnAsp 146
 Qy 292 TTTTCCCAACCTGGGGAGAAACCGCAGGAGATGGTTTGGCAAGAGTTCCCCCAACAA 351
 Db 147 PheProGlnAlaTrpAlaGluThrGlyGlyLeuGlyArgAlaIysCysGlnAlaProIle 166
 Qy 352 GTTATTCAACTGAAGCCAGTGGCCACACAGTGTCACTGACAGACAGTACCCCTTGAGTAA 411
 Db 167 IleIleAspLeuIysProThrAlaValProValSerIleIysGlnItyrProMetSerLeu 186

Qy 412 GAAGCTCAAGAGGAATTCGGCCGCATGTCCAAAGATTAAATCCACAGGCGCATCTAGTT 471
 Db 187 GluAlaHisMetGlyIleArgGlnHisIleIleIysPheLeuGluLeuGlyValLeuArg 206
 Qy 472 CTTGTCACATCTCCCTGGAATATCTCCCTGCTACCGGTTAGAAAGCTGGAGCTAATGAC 531
 Db 207 ProCysArgSerProTrpAsnThrProLeuLeuProValIysIysProGlyThrGlnAsp 226
 Qy 532 TATCGACCACTACAGGACTTGAGAGGTCAATAAACGGGTGCGAGGATATACACCCACA 591
 Db 227 TyrArgProValGlnAspLeuArgIleAsnIysArgThrValAspIleHisProThr 246
 Qy 592 GTCCCGAACCTTATACCTCTTGTGTGCTCTCCACCCCAACCGAGCTGTGTATACAGTA 651
 Db 247 ValProAsnProTyrAsnLeuLeuSerThrLeuIysProAspTyrSerTrpTyrThrVal 266
 Qy 652 TTGACTTAAAGATGCTCTTCTGCTGCTGAGATTACACCCCTACGCCAACCACTTTT 711
 Db 267 LeuAspLeuIysAspAlaPhePheCysLeuProLeuAlaProGlnSerGlnGluPhe 286
 Qy 712 GCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGCGAGCTCACCTCGACCCGACTG 771
 Db 287 AlaPheGluTrpIysAspProGluArgGlyIleSerGlyGlnLeuThrTrpThrArgLeu 306
 Qy 772 CCCCAAGGGTTCAAGAACTCCCGACCATCTTTGACGAAGCCCTTACACAGAGACTCGGC 831
 Db 307 ProGlnGlyPheIysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuThr 326
 Qy 832 AACTTCAGATCCAAACACCTCAGGTGACCTCTCTCCAGTACGTGGATACCTGCTTCGTG 891
 Db 327 AspPheArgThrGlnHisProGluValThrLeuLeuGlnItyrValAspAspLeuLeuLeu 346
 Qy 892 GCGGGAGCCACCAACAGGACTGCTTAGAAGGACCAAGGACACTACTGCTGGAAATCTCT 951
 Db 347 AlaAlaProThrIysIysAlaCysThrGlnGlyThrArgHisLeuLeuGlnGluLeuGly 366
 Qy 952 GACCTAGGCTACAGAGCTCTCTGTAAGAGCCCAAGATTTGCGAGGAGAGAGGTAAACATAC 1011
 Db 367 GluIysGlyTyrArgAlaSerAlaIysIysAlaGlnIleCysGlnThrIysValThrTyr 386
 Qy 1012 TTGGGTACAGTTTGGGGACGGGACGGATGCTGACGGAGGACGGAAGAAAGTGTGA 1071
 Db 387 LeuGlyTyrIleLeuSerGluGlyIysArgTrpLeuThrProGlyArgIleGluThrVal 406
 Qy 1072 GTCCAGATACCGCCCAACACAGCCCAACAAATGAGAGAGTTTGGGGACAGCTGGA 1131
 Db 407 AlaArgIleProProProArgAsnProArgGluValArgGluPheLeuGlyThrAlaGly 426
 Qy 1132 TTTTGCAGACTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAAC 1191
 Db 427 PheCysArgLeuTrpIleProGlyPheAlaGluLeuAlaAlaProLeuTyrAlaLeuThr 446
 Qy 1192 AAGAAAAAGGGAAATTCCTCGGCTCTCGACACAGAGGAGCATTTGATGCTATCAAA 1251
 Db 447 LysGluSerThrProPheThrTrpGlnThrGluHisGlnLeuAlaPheGluAlaLeuLys 466
 Qy 1252 AAGGCCCTGAGCGACCTGCTGCGCCCTCCCTGAGCTAACTAAACCTTACCTTACCTT 1311
 Db 467 LysAlaLeuLeuSerAlaProAlaLeuGlyLeuProAspThrSerIysProPheThrLeu 486
 Qy 1312 TATGTGATGAGCGTAAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGACCATGG 1371
 Db 487 PheLeuAspGluArgGlnGlyIleAlaLysGlyValLeuThrGlnIysLeuGlyProTrp 506
 Qy 1372 AGAAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1431
 Db 507 LysArgProValAlaTyrLeuSerIysLysLeuAspProValAlaAlaGlyTrpProPro 526
 Qy 1432 TGCCTGAAGGCTATCCAGCTGTGGCCATCTGCTCAAGGACGCTGACAAATGACTTTC 1491
 Db 527 CysLeuArgIleMetAlaAlaThrAlaMetLeuValIysAspSerAlaLysLeuThrLeu 546

A:Residues: 1-1199 <SHI>
 A:Cross-references: UNIPARC:UPI0000174A30
 A:Experimental source: clone pMLV-1
 A:Note: the pol polyprotein contains reverse transcriptase (about 80,000 daltons) and pol yet been defined
 C:Comment: This protein is synthesized as a gag-pol polyprotein.
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polyprotein
 C:Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polyprotein; reverse P;3-102/Product: retropepsin #status predicted <RTP>
 P;27/Active site: Asp (shared with dimeric partner) #status predicted

Alignment Scores:
 Pred. No.: 4.69e-246 Length: 1199
 Score: 3797.50 Matches: 719
 Percent Similarity: 76.3% Conservative: 165
 Best Local Similarity: 62.0% Mismatches: 242
 Query Match: 61.4% Indels: 33
 DB: 1 Gaps: 12

US-10-723-552-3_copy_2307_5741 (1-3435) x GNMV1M (1-1199)

```

QY 4 GGTGCCACGGGCAACACAGTATCCATGGACTACCCGAGAACAGTGTGAGTGGAGTG 63
DB 51 GlyAlaThrGlyGlyLysArgTyrArgTrpThrAspArgLysValHisLeuAlaThr 70
QY 64 GGCGGTGAACCACTCTGTTCTCGTCATACCTCAGTGCCCGACGACCCCTCTTAGGTAGA 123
DB 71 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 90
QY 124 GACTTATTGACCAAGATGGGAGCACAAATTTCTTTGAA---CAAGGGAAACACAGAAGTG 180
DB 91 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValMet 110
QY 181 TCTGCAATTAACAACCTATCAGTGTGTGACCTCCAAATAGATGAGCAATATCGACTA 240
DB 111 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu 130
QY 241 TACTCTCCCTAGTAAAGCCGTGATCAAAATATA---CAATTCCTGTTGGAAAGTATCCC 297
DB 131 HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 150
QY 298 CAAGCTGGCGCAGAAACCGCAGGATGGGTTGGCAAGCAAGATTCCTCCACCAAGTTATT 357
DB 151 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 170
QY 358 CAACTGAAGCCAGTGCACACCGAGTGTGATGACAGTACCGTCCCTTGAGTAAGAAGCT 417
DB 171 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrProMetSerGlnGluAla 190
QY 418 CAAGAAGGAATTGGCCGCATGTCCAAAGATTATCAACAGCGCATCTCTAGTTCCTGTC 477
DB 191 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 210
QY 478 CAATCTCCCTGGAACTATCCCTCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGA 537
DB 211 GlnSerProIleThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 230
QY 538 CCAGTACAGACTTGAGAGAGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCG 597
DB 231 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 250
QY 598 AACCTTATACCTCTGTGTGCTCTCCACCCCGACGAGCTGGTATACAGTATTGGAC 657
DB 251 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp 270
QY 658 TTAAGGATGCTTCTCTGCTGAGATTACACCCCACTAGCAACCACTTTTGGCTTC 717
DB 271 LeuLysAspAlaPhePheCysLeuArgGluHisIleProThrSerGlnProLeuPheAlaPhe 290
QY 718 GAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACTCTGGACCCCGACTGCCCAA 777
DB 718 GAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACTCTGGACCCCGACTGCCCAA 777

```

```

DB 291 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 310
QY 778 GGGTTCAAGAACTCCCGACCATCTTTGACGAAGCCCTTACACAGAGACTCGCCCAACTTC 837
DB 311 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 330
QY 838 AGGATCCAAACACCTCAGGTGACCCCTCTCCAGTACGTGGATGACCTGCTCTCGCGGA 897
DB 331 ArgIleGlnHisProAspLeuLeuLeuGlnTyrValAspAspLeuLeuAlaAla 350
QY 898 GCCACCAACAGGACTCTTAGAAGGCACGAGCAGTACTGCTGGAAATTTCTGACCTA 957
DB 351 ThrSerGlnLeuAspCysGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu 370
QY 958 GGCTACAGAGCTCTGCTAAGAGGCCCCAGATTTGACGAGGAGAGGTAACTACTTGGGG 1017
DB 371 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 390
QY 1018 TACAGTTTGGGACGGGACGATGCTGACGGAGCAGCGAAGAAACTGTAGTCCAG 1077
DB 391 TyrLeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly 410
QY 1078 ATACCGGCCCAACACACAGCACAACAAATACAGAGAGTTTGGGACAGCTGGATTTGC 1137
DB 411 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 430
QY 1138 AGACTGTGGATCCCGGGTTTGGACCTTAGCAGCCCTACTACCGCTACCAACCAAGAA 1197
DB 431 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 450
QY 1198 AAAGGGGAATTCCTCTGGGCTCTGAGCACACGAGGATTTGATGCTATCAAAAGGCC 1257
DB 451 GlyThrLeuPheAsnTrpGlyProAspGlnLysAlaTyrGlnGluIleLysGlnAla 470
QY 1258 CTGCTGAGCCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTACCTTTATGTG 1317
DB 471 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 490
QY 1318 GATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGCACCATGAGAGA 1377
DB 491 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArg 510
QY 1378 CCTGTCGCTACCTGTCAAAGAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGCTGT 1437
DB 511 ProValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTyrProCysLeu 530
QY 1438 AAGCTATCCAGCTGTGGCCATCTGCTCAAGACGCTACAAATTTGACTTTGGGACAG 1497
DB 531 ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln 550
QY 1498 AATATACTGTATAAGCCCTCATGATTGGAGAACATCGTTCCGACAGCCCGACAGCCGA 1557
DB 551 ProLeuValIleLeuAlaProHisAlaValGluAlaLeuValLysGlnProAspArg 570
QY 1558 TGGATGACCAACCCCGCATGACCCACTATCAAGGCTGCTCTCTC---ACAGAGAGGCTC 1614
DB 571 TrpLeuSerAsnAlaArgMetThrHisTyrGlnAlaLeuLeuLeuAspThrAspArgVal 590
QY 1615 ACCTTCGCTCCACGACCGCTCTCAACCCCTGCGCATCTTCTGCTGAGAGACTGATGAA 1674
DB 591 GlnPheGlyProValValAlaLeuAsnProAlaThrLeuLeuPro---LeuProGluGlu 609
QY 1675 CCAGTGACTCATGATTTGCCATCAACTATTGATTGAGGAGACTGGGTCCCGACAGGACCTT 1734
DB 610 GlyLeuGlnHisAsnCysLeuAspIleLeuAlaGluAlaHisGlyThrArgProAspLeu 629
QY 1735 ACAGATACATCCGCTGACTGGAGAAGTCTAACCTGGTTCACTGCGGAACAGCAGTATGTG 1794
DB 630 ThrAspGlnProLeuProAspAlaAspHisThrTrpTyrThrAspGlySerSerLeuLeu 649
QY 1795 GTGGAAGGTAAAGAGTGTGGCGCGGTGGTGGACGGGACCCGACGACTCTGGGCC 1854
DB 650 GlnGluGlyGlnArgLysAlaGlyAlaAlaValThrThrGluThrGluValIleTrpAla 669

```


790 LeuLeuIleGluAspSerThrProTyrThrProAlaTyrPheHisTyrThrGluThrAsp 809
2263 TGGCAAGAGATAAAAAAGATAGACACAGTTCTCTGAGACTCCGGAGGACCTGTATACC 2322
810 LeuLysIysLeuArgGluLeuGlyAlaThrTyrAsnGlnSerLysGly---TyrTrpVal 828
2323 TCAGATGGGAAGAAATCTGCCCCACAAAGAGGGTTAGAAATATGTCTCAACAGATACAT 2382
829 PheGlnGlyLysProValMetProAspGlnPheValPheGluLeuLeuAspSerLeuHis 848
2383 CGCTAACCCACCTAGGAACCTAAACACCTGCAGCAGATTGGTC-----AGAACATCCCT 2436
849 ArgLeuThrHisLeuGlyTyrGlnLysMetLysAlaLeuLeuAspArgGlyGluSerPro 868
2437 TATCATGTTCTGAGGCTACAGGAGTGCTGACTCGTGGTCAACATTGTGTCCTGCC 2496
869 TyrTyrMetLeuAsnArgAspLysThrLeuGlnTyrValAlaAspSerCysThrValCys 888
2497 CAGCTGGTTAATGCTAATCTTCCAGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCAC 2556
889 AlaGlnValAsnAlaSerLysAlaLysIleGlyAlaGlyValArgValArgGlyHisArg 908
2557 CCAGCGCTACTCGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAACAA 2616
909 ProGlyThrHisTrpGluIleAspPheThrGluValLysProGlyLeuTyrGlyTyrLys 928
2617 TACCTATTGGTTTGTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCTCTACTAAGAA 2676
929 TyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaPheProThrLysHis 948
2677 GAGACTTCAACCGTGGTGGCTAAATAAATACTGGAAGAAATTTTCCAAAGATTGGGAATA 2736
949 GluThrAlaLysIleValThrLysLysLeuLeuGluLullePheProArgPheGlyMet 968
2737 CTTAAGGTAAVAGGTCACAGATGGTCCAGCTTTTGTGGCCAGGTAAAGTCAGGACTG 2796
969 ProGlnValLeuGlyThrAspAsnGlyProAlaPheValSerGlnValSerGlnSerVal 988
2797 GCCAAGATATTGGGATTTGATTGGAACCTGATGTGTACATACAGACCCCAAGCTCAGGA 2856
989 AlaLysLeuLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSerGly 1008
2857 CAGGTAGAGAGATGAATAGAACCATTAAGAGACCTTACTTAATTCACCGCGAGACT 2916
1009 GlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuLeuAlaThr 1028
2917 GCGGTTAATGATTGGATAGCTCTCTCGCCCTTTTGTGCTTTTGTAGGGTTAGAACACCCCT 2976
1029 GlyThrArgAspTrpValLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThrPro 1048
2977 GGACAGTTTGGCTGACCCCTATGAATTAATCTACGGGGGACCCCGCCCATTTGGTAGAA 3036
1049 GlyProHisGlyLeuThrProTyrGluIleLeuTyrGlyAlaProProLeuLeuValAsn 1068
3037 ATTGCTCTGTACATAGTGTGACGTGCTGCTTTCCAGCCCTTTGCTCTAGGCTCAAG 3096
1069 PheHisAspProGluMetSerLysPheThrAsnSerProSerLeuGlnAlaHisLeuGln 1088
3097 GCATTTAGTGGGTGAGACAAACGACGCGTGGAGGCAACTCCGGAGGCGCTACTCAGGAGGA 3156
1089 AlaLeuGlnAlaValGlnArgGluValTrpLysProLeuAlaAlaAlaTyrGlnAspGln 1108
3157 GGAGACTTGCAG---ATCCACATCGTTTCCAGTGGGAGATTCACTAGTCTAGTTAGAGCG 3213
1109 LeuAspGlnProValIleProHisProPheArgValGlyAspThrValTrpValArgArg 1128
3214 CACCGTCGAGAAACCTCGAGACTCGGTGGAGGGCCCTTATCTCGTACTTTTGAACACA 3273
1129 HisGlnThrLysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThrThr 1148
3274 CCAACGGCTGTGAAGTGCAGAGGAATCTCCACCTGGATCCATGCATCCACGTTAAACCG 3333
1149 ProThrAlaLeuLysValAspGlyIleSerAlaTrpIleHisAlaAlaHisValLysAla 1168

3334 GCG-----CCACCTCCGATTCCGGGTGGAAAGCCGAAAGACTGAAAATCCCTTAAG 3387
1169 AlaThrThrProPheIleArgProSerTrpArgValGlnArgSerGlnAsnProLeuLys 1188
3388 CTTCGCTCCATCCGCTGGTTCCT 3411
1189 IleArgLeuThrArgGlyAlaPro 1196
RESULT 8
T10532
gag:pol polyprotein precursor - feline leukemia virus (strain FeLV-FAIDS)
N:Contains: core protein p10; core protein p12; core protein p15; core protein p27; endo
C:Species: feline leukemia virus
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10532
J:Donahue, P.R.; Hoover, E.A.; Beltz, G.A.; Riedel, N.; Hirsch, V.M.; Overbaugh, J.; Mul
R. Virol. 62, 722-731, 1988
A:Title: Strong sequence conservation among horizontally transmissible, minimally pathog
A:Reference number: Z17078; MUID:88119207; PMID:2828667
A:Accession: T10532
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1784 <DON>
A:Cross-references: UNIPROT:Q85521; UNIPARC:UPI000010C1F5; EMBL:M18247; NID:G323904; PID
C:Superfamily: pol polyprotein
C:Keywords: hydrolase; nucleotidyltransferase; polyprotein; reverse transcriptase
F:75-576/Product: gag polyprotein #status predicted <GAG>
F:75-201/Product: core protein p15 #status predicted <p15>
F:202-271/Product: core protein p12 #status predicted <p12>
F:272-519/Product: core protein p27 #status predicted <p27>
F:520-576/Product: core protein p10 #status predicted <p10>
F:577-1784/Product: pol polyprotein #status predicted <POL>
F:577-701/Product: proteinase #status predicted <PIN>
F:702-1368/Product: RNA-directed DNA polymerase (EC 2.7.7.49) #status predicted <REV>
F:1369-1784/Product: endonuclease (EC 3.1.-.-) #status predicted <EDC>
Alignment Scores:
Pred. No.: 3,25e-239 Length: 1784
Score: 3696.00 Matches: 706
Percent Similarity: 74.3% Conservative: 158
Best Local Similarity: 60.7% Mismatches: 257
Query Match: 59.8% Indels: 42
DB: 2 Gaps: 12
US-10-723-552-3_copy_2307_5741 (1-3435) x T10532 (1-1784)
QY 4 GGTGCCACAGGCAACACAGTATCCATGGACTACCCGAGAACAGTTGACTTGGAGTG 63
DB 632 GlyAlaThrGlySerLysAsnTyrArgTrpThrThrAspArgValGlnLeuAlaThr 651
QY 64 GGACGGTAAACCACTCGTTTCTGGTCATACCTGAGTCCAGCACCCCTCTTAGGTAGA 123
DB 652 GlyLysValThrHisSerPheLeuTyrValProGluCysProTyrProLeuLeuGlyArg 671
QY 124 GACTTATTGACCAAGATGGGAGACAAATTTCTTTTGAA---CAAGGGAACACGAAAGTG 180
DB 672 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheThrGlyGluGlyAlaAsnValVal 691
QY 181 TCTGCAATATCAAAACCTTATCATCTGTGTGACCTCCCAATTAGATGACGAATATCGACTA 240
DB 692 GlyProArgGlyLeuProLeuGlnValLeuThrLeuGlnLeuGluGluTyrArgLeu 711
QY 241 TACTCTCCCTAGTAAGCCCTGATCAAAATATACAAATTCCTGTTGGAAACAGTTCCTCCCAA 300
DB 712 PheGluProGluSerThrGlnLysGlnGluMetAspIleTrpLeuLysAsnPheProGln 731
QY 301 GCCTGGGACAGAAACCGCAGGGATGGTTTGGCAAGCAAGTTCCTCCCAACAGTTATTCAA 360
DB 732 AlaTrpAlaGluThrGlyGlyMetAlaHisCysGlnAlaProValLeuIleGln 751
QY 361 CTGAAGGCCAGTGCACACCAAGTGTCAAGTACAGACAGTACCCCTTTCAGTAAAGAGCTCAA 420

Db 241 ValThrTyrLeuGlyTyrIleLeuSerGluGlyLysArgTrpLeuThrProGlyArgIle 260
Qy 1063 AAAAAGTGTAGTCCAGATACCGCCGCCCAACACACAGCCAAACAAATGAGAGAGTGTTCGGG 1122
Db 261 GluThrValAlaHisIleProProGlnAsnProArgGluValArgGluPheLeuGly 280
Qy 1123 ACAGCTGGATTTTCAGACTGTGATCCCGGGTTTGGACCTTAGCAGCCCACTCTAC 1182
Db 281 ThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGluLeuAlaAProLeuTyr 300
Qy 1183 CCGCTAACCAAGAAAAGGAATTCCTGGGCTCTGAGCACCAAGGAGCATTTGAT 1242
Db 301 AlaLeuThrLysGluSerAlaProPheThrTrpGlnGluLysHisGlnSerAlaPheGlu 320
Qy 1243 GCTATCAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCC 1302
Db 321 AlaLeuLysGluAlaLeuLeuSerAlaProAlaLeuGlyLeuProAspThrSerLysPro 340
Qy 1303 TTTACCTTTATGTGGATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACCAACCCCTA 1362
Db 341 PheThrLeuPheIleAspGluLysGlnGlyIleAlaLysGlyValLeuThrGlnLysLeu 360
Qy 1363 GGACCATGGAGAACCTGTGCTGCTACCTGTCAAGAGAGCTCGATCTGTAGCCAGTGT 1422
Db 361 GlyProTrpLysArgProValAlaTyrLeuSerLysLysLeuAspProValAlaIleArg 380
Qy 1423 TGGCCCATATGCTGAAGGCTATCGAGCTGTGGCCATATCTGGTCAAGGACGCTGACAAA 1482
Db 381 TrpProProCysLeuAArgIleMetAlaAlaThrAlaMetLeuValLysAspSerAlaLys 400
Qy 1483 TTGACTTTGGGACAGATATAACTGTATATAGCCCCCATGATGGAGAGACATGTTCCG 1542
Db 401 LeuThrLeuGlyGlnProLeuThrValIleThrProHisAlaLeuGluAlaIleValArg 420
Qy 1543 CAGCCCCGACGACGATGACCAACCGCGCATGACCCACTATCAAGCGCTCTCTC 1602
Db 421 GlnThrProAspArgTrpIleThrAsnAlaArgLeuThrHisTyrGlnAlaLeuLeuLeu 440
Qy 1603 ---ACAGAGGGTTCAGTTCGCTCCACACGCGCTCTCAACCTGCGCACTCTCTCGCT 1659
Db 441 AspThrAspArgIleGlnPheGlyProProValThrLeuAsnProAlaThrLeuLeuPro 460
Qy 1660 GAAGAGACTGATGAACCAAGTACTCATGATGGCATCACTATTGATGGAGAGACTGG 1719
Db 461 AlaProGluAspGlnGlnSerAlaHisAspCysArgGlnValLeuAlaGluThrHisGly 480
Qy 1720 GTCGCAAGACCTTACAGACATACCGCTGACTGGAGAGTGTAACTGGTCACTGAC 1779
Db 481 ThrArgGluAspLeuLysAspGlnGluLeuProAspAlaAspHisSerTrpTyrThrAsp 500
Qy 1780 GGAAGCAGCTATGTGGTGAAGGTAAAGAGATGGCTGGCGCGGTGGTGGACGGGACC 1839
Db 501 GlySerSerTyrIleAspSerGlyThrArgArgAlaGlyAlaValValAspGlyHis 520
Qy 1840 CGCAGCATCTGGCGACAGCCTGCCGGAAGGAATTCAGACAAAAGGTGAGCTCATG 1899
Db 521 HisIleIleTrpAlaGlnSerLeuProGlyThrSerAlaGlnLysAlaGluLeuIle 540
Qy 1900 GCCTCAGCAAGCTTTGCGGCTGGCGAAGGGAAATCCATAACATTTATACGGACAGC 1959
Db 541 AlaLeuThrLysAlaLeuGluLeuSerGluGlyLysAlaAsnIleTyrThrAspSer 560
Qy 1960 AGGTATGCTTTGCGACTGCACAGTACATGCGGCGCATCTATAAAGGGGGTTCCTT 2019
Db 561 ArgTyrAlaPheAlaThrAlaHisThrHisGlySerIleTyrGluArgArgGlyLeuLeu 580
Qy 2020 ACCTCAGCAGGGAGGGAATAAAGAACAAAGAGAGAAATTCACCCCTATTAGAACCGTA 2079
Db 581 ThrSerGluGlyLysGluIleLysAsnLysAlaGluIleIleAlaLeuLeuLysAlaLeu 600
Qy 2080 CATTTACCAAAAAGGCTAGCTATTATACATGCTGCGACATCAGAAAGCTAAAGATCTC 2139
Db 601 PheLeuProArgLysValAlaIleIleHisCysProGlyHisGlnLysGlyGlnAspPro 620

Qy 2140 ATATCCAGAGAAACACAGATGGCTGACCGGGTTCCCAAGCAG-----GCAGCCACGGGT 2193
Db 621 IleAlaThrGlyAsnArgGlnAlaAspGlnValAlaArgGlnValAlaValAlaGluThr 640
Qy 2194 GTTAACCTTTCGCTATATAATAGAAATGCCCAAGCCCCAGAACCCAGACGACGATACAC 2253
Db 641 LeuThrLeuThrThrLysLeuGluGluThrAsnLeuThrThrAsnLysTyrAlaTyrThr 660
Qy 2254 CTGAAGACATGGCAAGAGATAAANAACATAGACACCGATTCT-----TCTGAGACTCCGAA 2307
Db 661 ProGluAspGlnGluGluAlaLysAlaIleGlyAlaIleLeuAsnGlnAspThrLysAsp 680
Qy 2308 GGGACCTGCTATACCTCAGATGGGAAGAAATCTCGCCCCCACAAGAGGGTGAAGATAT 2367
Db 681 -----TrpGluLysGluGlyLysIleValLeuProArgLysGluAlaLeuAlaMet 697
Qy 2368 GTCACACAGATACATCTCTTAACCCACTAGGAACCTAAACACCTGACGAGTGTGTCAGA 2427
Db 698 IleGlnGlnMetHisAlaTrpThrHisLeuSerAsnGlnLysLeuLysLeuLeuIleGlu 717
Qy 2428 ACATCCCTTATCATGTTCTGAGGCTACACGAGGTGCTGACTCGGTGGTCAACATGTT 2487
Db 718 LysThrAspPheLeuIleProLysAlaGlyThrLeuIleGluGlnValThrSerAlaCys 737
Qy 2488 GTGCCCTGCCAGCTGGTTAATGCTTAATCTTCCAGAATGCTTCCAGGGAAGAGACTAAGG 2547
Db 738 LysValCysGlnGlnValAsnAlaGlyAlaThrArgValProGluGlyLysArgThrArg 757
Qy 2548 GGAAGCCACCCAGCGCTCACTCGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAATATAC 2607
Db 758 GlyAsnArgProGlyValTyrTrpGluIleAspPheThrGluValLysProHisTyrAla 777
Qy 2608 GGAACAAATACCTATTGTTTGTAGACACCTTTTCCAGATGGGTAGAGGCTTATCCT 2667
Db 778 GlyTyrLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyrPro 797
Qy 2668 ACTAAGAAAGAGACTTCAACCGTGGTGGCTTAAATAAATACTTGAAGAAATTTTCCAAAGA 2727
Db 798 ThrArgGlnGluThrAlaHisMetValAlaLysLysIleLeuGluGluIlePheProArg 817
Qy 2728 TTTGGAATACCTAAGGTAAATAGGTGACAAATGGTCCAGCTTTTGTGGCCAGGTAACT 2787
Db 818 PheGlyLeuProLysValIleGlySerAspAsnGlyProAlaPheValSerGlnValSer 837
Qy 2788 CAGGACTGCGCAAGATATTTGGGATTTGATTGGAACCTGCTGTCATACAGACCCCA 2847
Db 838 GlnGlyLeuAlaArgThrLeuGlyIleAsnTrpLysLeuHisCysAlaTyrArgProGln 857
Qy 2848 AGCTCAGGACAGGTAGAGAGATCAATAGAACCATTAAGAGAGACCTTACTAAATTGACC 2907
Db 858 SerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThr 877
Qy 2908 CGCGAGACTGCGCTTAAATGATGATGATGCTCTCTGCGCTTGTGCTTTTGGGTAGG 2967
Db 878 LeuGluThrGlyLeuLysAspTrpArgArgLeuLeuSerLeuAlaLeuLeuAlaArg 897
Qy 2968 AACACCCCTGACAGTTTGGCTGACCCCTATGAATTAATCTACTACGGGGGACCCCCCA 3027
Db 898 AsnThrProAsnArgPheGlyLeuThrProTyrGluIleLeuTyrGlyGlyProProPro 917
Qy 3028 TTGCTGAATAATCTCTGTACATAGTGTGCTGAGCTGCTGCTTCCAGCCT----- 3078
Db 918 LeuSerThrLeuLeuAsnSerPheSerPro-----SerAspProLysThrAsp 933
Qy 3079 TTCTTCTTAGGCTCAAGGCACCTTGATGGTGGTGAGACAAACGAGCTGGAGGCAACTCCGG 3138
Db 934 LeuGlnAlaArgLeuLysGlyLeuGlnAlaValGlnAlaGlnIleTrpThrProLeuAla 953
Qy 3139 GAGGCTTACTCAGGAGGAGACTTGCAGATCCCATCGTTTCCAGTGGGAGATTC 3198
Db 954 GluLeuTyrArgProGlyHisPro---GlnThrSerTyrProPheGlnValGlyAspSer 972

Db 201 LysAspGlnProLeuProGly---ValProAlaTrpTyrThrAspGlySerSerPheIle 219

Qy 1795 GTGGAGGTAAAGATGCTGGGGCGCGGTGTGGACGGGACCCGACGATCTGGGCC 1854

Db 220 AlaGluGlyLysArgAlaGlyAlaAlaIleValAspGlyLysArgThrValTrpAla 239

Qy 1855 AGCAGCTTCGCGGAGGAACCTCAGCAAAAGGCTGAGCTCATGGCCCTCAGCCAGCT 1914

Db 240 SerSerLeuProGluGlyThrSerAlaGlnLysAlaGluLeuValAlaLeuThrGlnAla 259

Qy 1915 TTGGCGCTGGCGGAGGAATCCATAAACATTTATACGACAGCAGGTATGCTTTGCG 1974

Db 260 LeuArgLeuAlaGluGlyArgAspIleAsnIleTyrThrAspSerArgTyrAlaPheAla 279

Qy 1975 ACTGCACGTTACATGGGGCCATCTATAAACAAGGGGTGCTTACCTCAGCAGGGAGG 2034

Db 280 ThrAlaHisIleHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyLys 299

Qy 2035 GAAATAAAGAACAAAGAGGAATCTTAAGCCTATTAGAAGCGGTACATTTACCAAAAGG 2094

Db 300 AspIleLysAsnLysGluGluLeuAlaLeuLeuGluAlaIleHisLeuProLysArg 319

Qy 2095 CTAGCTATTATACACTGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAC 2154

Db 320 ValAlaIleIleHisCysProGlyHisGlnLysGlyAsnAspProValAlaThrGlyAsn 339

Qy 2155 CAGATGGCTGACGGGTGCCAAGCAGCAGCCAGGGTGTAACTTCTCGCTTATATA 2214

Db 340 ArgArgAlaAspGluAlaLysGlnAlaLeuSerThrArgValLeuAlaGluThr 359

Qy 2215 GAAATGCCAAGCCCAAGCCAGACCCAGACGACAGTACACCTTAGAAGCTGCCAGAGATA 2274

Db 360 ThrLysPro-----GlnGluLeu 365

Qy 2275 AAAAAGATAGACCACTTCTGAGACTCCGGAAGGGACCTGCTATACCTCAGATGGGAAG 2334

Db 366 IleLysProAlaGlnValLysThrArgPro-----Gly 376

Qy 2335 GAAATCTGCCCCCAAGAAGGGTTAGATATGTCACAGATACATCGTCTAACCCAC 2394

Db 377 GluLeuThrProAspArg--GlyLysGluPheIleGlnArgLeuHisGlnLeuThrHis 395

Qy 2395 CTAGGAACCTAACACCTGCAGCAGTTGCTCAGAACATCCCTTATCATGTTCTGAGGCTA 2454

Db 396 LeuGlyProGluLysLeuLeuGlnLeuValAsnArgThrSerLeuLeuIleProAsnLeu 415

Qy 2455 CCAGAGTGGCTGACTCGCGTGGTCAAAATTTGTGTGTCCTCCAGCTGGTTAAATGCTAAT 2514

Db 416 GlnSerAlaValArgGluValThrSerGlnCysGlnAlaCysAlaMetThrAsnAlaVal 435

Qy 2515 CTTCCAGAAATGCTCCAGGAGAGACTAAGGGGAAGCCACCCAGGCGCTCAGTGGGAA 2574

Db 436 ThrThrTyrArgGluThrGlyLysArgGlnArgGlyAspArgProGlyValTyrTrpGlu 455

Qy 2575 GTGACCTTCACTGAGTAAAGCCGCTAAATACGGAACAAATACCTATTGTTTTCGTA 2634

Db 456 ValAspPheThrGluValLysProGlyArgTyrGlyAsnArgTyrLeuLeuValPheIle 475

Qy 2635 GAC 2637

Db 476 Asp 476

RESULT 12

A42743

pol polyprotein - radiation murine leukemia virus (strain Kaplan) (fragment)

N:Contains: endonuclease (EC 3.1.-.-); proteinase (EC 3.4.21.-); RNA-directed DNA polym

C:Species: radiation murine leukemia virus

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004

C:Accession: A42743

R:Poliquin, L.; Bergeron, D.; Fortier, J.L.; Paquette, Y.; Bergeron, R.; Rassart, E.

J. Virol. 66, 5141-5146, 1992

A:Title: Determinants of thymotropism in Kaplan radiation leukemia virus and nucleotide

A:Reference number: A42743; MUID:92333703; PMID:1629969

A:Accession: A42743

A:Molecule type: DNA

A:Residues: 1-581 <POL>

A:Cross-references: UNIPROT:P31795; UNIPARC:UPI0000131P09; GB:M93052; NID:g332065; PIDN:

C:Comment: This protein is likely to be expressed as a fused gag-pol polyprotein.

C:Genetics: The precise boundary between the mature proteins has not been determined.

C:Gene: pol

C:Superfamily: pol polyprotein

C:Keywords: endonuclease; hydrolase; nucleotidyltransferase; polyprotein; reverse trans

Alignment Scores:

Pred. No.: 1.64e-104 Length: 581

Score: 1681.00 Matches: 329

Percent Similarity: 69.9% Conservatives: 84

Best Local Similarity: 55.7% Mismatches: 148

Query Match: 27.2% Indels: 30

DB: Gaps: 6

US-10-723-552-3_COPY_2307_5741 (1-3435) x A42743 (1-581)

Qy 1696 CAACATTATTGATTGAGGAGACTGGGTCCGCAAGACCTTACAGACATACCCGTGACTGGA 1755

Db 2 GluIleLeuAlaGluThrHisGlyThrArgProAspLeuThrAspGlnProIleProAsp 21

Qy 1756 GAAGTGTCTAACCTGGTTCACCTGACGGAAGCAGCTATGTGTGGAGGTAAAGAGGATGCT 1815

Db 22 AlaAspHisThrTrpTyrThrAspGlySerSerPheLeuGlnGluGlyGlnArgLysAla 41

Qy 1816 GGGCGCGGTGTGGACGGGACCCGACGATCTGGGCGCAGCAGCTGCCGGAAGGAACT 1875

Db 42 GlyAlaAlaValThrThrGluThrGluValIleTrpAlaArgAlaLeuProAlaGlyThr 61

Qy 1876 TCAGACAAAAAGGTGAGCTCATCGCCTCAGCAAGCTTTGCGGTGGCGAGGGGAAA 1935

Db 62 SerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLys 81

Qy 1936 TCCATAACATTTATACGACAGCAGGTATGCTTGGCTGCGACATGATCGGCTGCGGCC 1995

Db 82 ArgLeuAsnValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyGlu 101

Qy 1996 ATCTATAACAAAGGGGTGCTTACCTCAGCAGGAGGGAATAAAGAAACAAAGAGGAA 2055

Db 102 IleTyrLysArgArgGlyLeuLeuThrSerGluArgGluIleLysAsnLysSerGlu 121

Qy 2056 ATTCTAAGCTTATTAGAGCGGTACATTACCAAAAAGGTAGCTTATTATACATGTCCT 2115

Db 122 IleLeuAlaLeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysLeu 141

Qy 2116 GGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAACAGATGGCTGACCGGTTGCC 2175

Db 142 GlyHisGlnLysGlyAspSerAlaGluAlaArgGlyAsnArgLeuAlaAspGlnAlaAla 161

Qy 2176 AAGCAGGCGAGCCAGGGGTGTTAACTTCTGCTTATAAGAAATGCCCAAGCCCCAGAA 2235

Db 162 ArgGluAlaAla-----IleLysAlaProPro 170

Qy 2236 CCCAGACAGTATACCCCTAGAGAC----- 2262

Db 171 AspThrSerThrLeuLeuIleGluAspSerThrProTyrThrProAlaTyrPheHisTyr 190

Qy 2263 ---TGGCAAGAGATAAAAAAGATAGACAGTTCTCTGAGACTCCGGAAGGAGCC----- 2313

Db 191 ThrGluThrAspLeuLysLysLysLeuArgGluLeuGlyAlaThrTyrAsnGlnSerGlnGly 210

Qy 2314 TGCTATACCTCAGATGGGAAGAAATCTGCCCCACAAAGAGGGTTAGATAATGTGCCAA 2373

Db 211 TyrTrpValPheGlnGlyLysProValMetProAspGlnPheValPheGluLeuLeuAsp 230

Qy 2374 CAGATACATCTCTTAACCCACTAGGAACTAAACACCTGCAGCAGTGGTGGTC-----AGA 2427

Db 231 SerLeuHisArgLeuThrHisLeuGlyTyrGlnLysMetLysAlaLeuLeuAspArgGly 250

RESULT 15
B46312
pol polyp

C;Species: human endogenous virus S71
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: B46312
R;Werner, T.; Brack-Werner, R.; Leib-Moesch, C.; Backhaus, H.; Erfle, V.; Hehlmann, R.
Virology 174, 225-238, 1990
A;Title: S71 is a phylogenetically distinct human endogenous retroviral element with structural similarity to HTLV-1
A;Reference number: B46312; MUID:90101379; PMID:2152993
A;Accession: B46312
A;Molecule type: DNA
A;Residues: 1-602 <WER>
A;Cross-references: UNIPROT:Q7LYY5; UNIPARC:UPI000017A87B; GB:M32788
C;Genetics:
A;Gene: pol
C;Keywords: polyprotein

Alignment Scores:
Pred. No.: 1,79e-68 Length: 602
Score: 1142.50 Matches: 283
Percent Similarity: 57.1% Conservative: 75
Best Local Similarity: 45.1% Mismatches: 203
Query Match: 18.5% Indels: 67
DB: 2 Gaps: 18

US-10-723-552-3_COPY_2307_5741 (1-3435) x B46312 (1-602)

Qy	1378	CTGTGCGCTACCTGTCGAAGAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGCTG	1437
Db	2	ProValIleTyrLeuSerLysLeuAspProValAlaSerArgTrpThrSerCysLeu	21
Qy	1438	AAGGCTATCGGAGCTGTGGCCATCTGTCGAAGAGCGCTGCAAAATTCACCTTTGGGACAG	1497
Db	22	TrpAlaIleAlaAlaThrAlaSerLeuIleGlnThrAspLysLeuThrLeuSerGln	41
Qy	1498	AATTAACCTGAATAGCCCCCATGTCATGAGAACATCGTTGGCAGCCGCCACGACGGA	1557
Db	42	AsnLeuThrLeuAlaProArgAlaIleGlnThrLeuLeuGlnSerAlaSerGlyLys	61
Qy	1558	TGGATGACCAAGCCCGCATGACCCATCATTAAGCCCTGCTTCACACA--GAGAGGCTC	1614
Db	62	TrpMetSerAsnAlaArgIleLeuGlnTyrGlnSerLeuLeuLeuAspTrpProArgLeu	81
Qy	1615	ACGTTGCTGCCACGCGCTCTCAACCTGCGCATCTTCTGCTGGAAGAGACGATGAA	1674
Db	82	ThrPheSerProThrArgCysLeuAsnProAlaThrLeuLeuProAspProAspPheThr	101
Qy	1675	CCAGTGACTCATGATGTCATCACTATTGATAGAGACCTGGGGTCCGCAAGACCTT	1734
Db	102	ThrProValHisAspCysGlnGluLeuGlnThrThrGluThr---ValArgProAspLeu	120
Qy	1735	ACAGACATACCGCTGACTGGAGAGTGTAACTGGTTCACTGACGGAGCAGCTATG	1794
Db	121	GlnAspValProLeuLysGluValAspAlaThrValPheThrAspSerSerLeuLeu	140
Qy	1795	GTGGAGGTGAAGAGGATCGTGGGGCGGGTGGTGGAGCGGACCCGACCATCTGGG	1854
Db	141	LysGlnGlyValArgLysAlaGlyAlaAlaValThrMetGluThrAspLysLeuGlnThr	160
Qy	1855	AGCAGCTGCCGGAAGAACTTCAGCAAAAGGCTGAGCTCATGGCCCTCAGCGAAGCT	1914
Db	161	GlnAlaLeuProAlaGlyThrSerAlaGlnLysAlaGluLeuValAlaLeuIleGlnAla	180
Qy	1915	TTGCGGCTGGCGAAGGGAATCCATAAACAATTTATACGGACAGCAGCTATGCC--TTTGC	1973
Db	181	LeuArgArgVal-ArgThrAsnValLeuThrPheThrLeuThrAlaGlyMetLeuPheAl	200
Qy	1974	GACTGCACATGATCGGGCCATCTATAAACAAGGGGGTGTCTTACTCTAGCAGGAG	2033
Db	200	aThrValArgValHisGlyAlaIleTyrGlnValArgGlyLeuLeuThrSerAlaGlyLys	220
Qy	2034	GGAATAAAGAACAAAGAGGAATTTCTAAGCTATTAGAAGCCGTACATTTCCAAAG	2093
Db	220	SalAileLysAsn---GluGluIleLeuAlaLeuLeuGluAlaValCysLeuProGlnG	239

Qy	2094	GCTAGCTATTATACACTGTCTCTGACATCAAAAGCTAAAGATCTCATATCCAGAGGAAA	2153
Db	239	nValAlaValIleHisCysLysGlyHisGlnLysGluAspThrAlaValAlaHisGlyAs	259
Qy	2154	CCAGATGGCTGACCGGGTTGCCAAGCAGCGCAGCCAG-----GGTTTAACTT	2201
Db	259	nGlnArgAlaAspSerAlaAlaTrpGlyProAlaGlnLeuProValAlaProThrLe	279
Qy	2202	TCGTCTATATAGAAATGCCAAAGCCCCCAGAACCCAGACGACGATACCCCTAGAGA	2261
Db	279	uLeuProAlaValSerPhePro-----GlnProAspLeuSerAs	292
Qy	2262	CTGGCAAGAGATAAAAGATAGACCACTTCTCTGAGACTCCGGAAGGAGCCTGCTATAC	2321
Db	292	pHisProGluTyr-----SerProGluGluGluLysGlnAl	304
Qy	2322	CTCAGATGGGAAGAAATCTGCCCCCAAAAGAGGGTTAGAAATATGTCCAACAGATACA	2381
Db	304	aSerAsp---LeuGlnAlaSerLysAsnGlnGluGly-----	315
Qy	2382	TCGTCTAACCCACCTAGGACTAAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCA	2441
Db	316	-----GlyValLys---LeuAlaGlnLeuLeuArg---SerArgPheLys	328
Qy	2442	TGTTCTGAGGCTACACGAGGTGGCTGACTCGTGGTCAAAACATTTGTGTGCTGCTGCAGCT	2501
Db	328	sIleProAsnLeuGlnAspLeuValAsnGlnAlaAlaLeuTrpCysThrValCysAlaG	348
Qy	2502	GGTTAATGCTAATCCTTCCAGAATGCCT---CCAGGGAAGAGACTAAAGGGAAGCCACC	2558
Db	348	nValAsnThrLysGlnGlyProLysProSerSerGlyAspArgLeuGlnGlyAspSerPr	368
Qy	2559	AGGCGCTCAGTGGGAAGTGGACTTCACCTGAGTAAAGCCGCTAAATACGGAACAATA	2618
Db	368	oGlyGluArgTrpGluLeu-----ThrGluIleLysProHisTrpAlaGlyTyrLysTy	386
Qy	2619	CCTATTGTTTCTGACACCTTTTCAGGATGGGTAGAGCTTATCTCTACTAAGAAAGA	2678
Db	386	rLeuLeuValLeuValAspThrPheSerGly---ThrGluAlaPheAlaThrLysAsnG	405
Qy	2679	GACTTCAACCTGTGTGCTAAAAAATACTGGAAGAAATTTTCCAGAAATTTGGAATACC	2738
Db	405	uThrAlaThrThrValValLysPheSerLeuAsnGluIleLeProGlnHisGlyLeuPr	425
Qy	2739	TAAGTGAATAGGTCAGACATGCTCCAGCTTTTGTGCCCCAGGTAACTAGTCCAGGACTGC	2798
Db	425	oThrAlaMetGlySerAspAsnArgSerAlaPheThrSerSerIleAlaGlnSerValSe	445
Qy	2799	CAAGATATTGGGATTGATTGAAACTGCATTGTGTCATACAGACCCCAAGCTCAGGACA	2858
Db	445	rLysAlaLeuAsnIleGlnTrpLysLeuArgCysAlaTyrArgProGlnSerSerGlyTr	465
Qy	2859	GGTAGAGGATGAATAGAACCATTAAGAGACCCCTTACTAAATTTGACCCGGGAGACTGG	2918
Db	465	pValGluHisMetAsnHisThrLeuLysAsnThrValThrLysLeuIleLeuGluThrGl	485
Qy	2919	CGTTAATGATGGATAGCTCTCTCCCTTTTGTGCTTTTAGTGGTAGACACCCCTGG	2978
Db	485	yLysAsnGln---ValArgLeuLeuProLeuThrLeuLeuLysValArgCysIleProTy	504
Qy	2979	ACAGTTTGGGCTGACCCCTATGAATTAATCTAGCGGGGACCCCTCCCATTTGGTA----	3033
Db	504	rArgAlaGlyPheSerProPheGluIleThrTyrArgArgArgProPheLeuLeuProLy	524
Qy	3034	-----GAAATTTGCTTGTACATAGTGTGACGTGCTGCTGCTTTTCCCGACCTTT	3080
Db	524	sLeuLysAspThrArgLeuAlaGluIleSerGluAlaAsnLeuLeu-----GlnTyrLe	542
Qy	3081	GTTCTCTAGGCTCAGGCACCTTGAGTGGTGAGACACAGCGGTGGAGGCACTCCGGGA	3140
Db	542	uGlnSerLeuGlnValArgAspIleIleGlnProLeuValTrp-----	557
Qy	3141	GGCCTTACTCAGGAGGAGGAGACTTTGCAGATCCCA-----CATCGTTT	3182

Db	558	-----GlyAlaHisProSerProValProAspGlnThrGlyProCysHisSerPh		:::		
Qy	3183	CCAAAGTGGAGATTCA				
Db	574	eProProGlyAspLeuVal				

Search completed: February 14, 2006, 16:06:40
Job time : 286.291 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.
OM nucleic - protein search, using frame_plus_n2p model
Run on: February 14, 2006, 12:52:14 ; Search time 142.918 Seconds
(without alignments)
3391.447 Million cell updates/sec
Title: US-10-723-552-3_COPY_2307_5741
Perfect score: 6183
Sequence: 1 ATGGTGCCACAGGCAACA.....CTGTCAATAACCTCTCAGAC 3435
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 4332886
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10723552/runat.14022006.125141.12876/app.query.fasta_1
-DB-UniProt -QFMT=fasta -SUFFIX=rup -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USR=US10723552 @CGN 1.1 808 @runat.14022006.125141.12876 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot.sprot.*
2: uniprot_trembl.*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Description
1	6026	97.5	1145	2	Q9XSN8_PIG
2	5922.5	96.9	1193	2	Q5QG8_PORCINE
3	5986	96.8	1145	2	Q4VFZ2_PORCINE
4	5914.5	95.7	1193	2	Q90RL9_PORCINE
5	5885	95.2	2378	2	Q9Q1X4_PORCINE
6	5872	95.0	1720	2	Q8J4V6_PORCINE
7	5866.5	94.9	1718	2	Q8J4V8_PORCINE
8	5864	94.8	2376	2	Q9Q1X5_PORCINE
9	5863	94.8	1146	2	Q8UMP5_PORCINE
10	5861	94.8	1195	2	Q90RL6_PORCINE
11	5861	94.8	2376	2	Q9Q1X3_PORCINE
12	5840.5	94.5	1144	2	Q8UM99_PORCINE
13	5831	94.3	1147	2	Q8UM96_PORCINE
14	5466.5	88.4	1142	2	Q8Q6U4_PORCINE
15	5383.5	87.1	1139	2	Q8Q6U7_PORCINE
16	4450	72.0	868	2	Q73505_PORCINE

17	4156.5	67.2	1127	2	Q9TTC1_PHACI
18	4146.5	67.1	1203	2	O89815_MUS
19	4138.5	66.9	1127	2	O70652_GALV
20	4135.5	66.9	1165	1	POL_GALV
21	3835	62.0	1734	2	O7SVK7_GAMR
22	3833	62.0	1196	1	POL_MLAV
23	3831	62.0	1733	2	Q9E7M1_GAMR
24	3825	61.9	1734	2	Q9J8E2_GAMR
25	3818.5	61.8	1204	1	POL_MLVP
26	3818	61.7	1196	2	Q90RL4_GAMR
27	3818	61.7	1734	2	P70355_MOUSE
28	3815.5	61.7	1199	2	Q60FS9_GAMR
29	3814.5	61.7	1736	2	Q7ZJT6_GAMR
30	3811.5	61.6	1204	2	O41250_MLVR
31	3810.5	61.6	1199	2	Q60FS6_GAMR
32	3810.5	61.6	1738	2	O39735_MLVR
33	3809.5	61.6	1204	1	POL_MLVP
34	3808.5	61.6	1204	2	Q9YK99_GAMR
35	3802.5	61.5	1189	1	POL_BAEYM
36	3797.5	61.4	1199	1	POL_MLVM
37	3797.5	61.4	1737	2	O92808_MLVM
38	3797.5	61.4	1738	2	Q8UN00_MLVM
39	3785.5	61.2	1204	1	POL_MLVP
40	3784.5	61.2	1736	2	Q5PYI3_GAMR
41	3778	61.1	1196	1	POL_MLVRD
42	3772	61.0	1204	2	Q7ZKZ7_GAMR
43	3757	60.8	1736	2	O83362_GAMR
44	3719	60.1	1786	2	O89811_FLV
45	3718.5	60.1	1180	2	Q9WHV7_GAMR

ALIGNMENTS

RESULT 1
Q9XSN8_PIG
ID Q9XSN8_PIG PRELIMINARY; PRT; 1145 AA.
AC Q9XSN8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol protein.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=miniature swine;
RX MEDLINE=98216827; PubMed=9557749;
RA Akiyoshi D.E., Denaro M., Zhu H., Greenstein J.L., Banerjee P.,
RA Fishman J.A.
RT "Identification of a full-length cDNA for an endogenous retrovirus of
RT miniature swine."
RL J. Virol. 72:4503-4507(1998).
DR EMBL; AF038600; AAC16764.1; -, mRNA.
DR HSSP; P03355; 116J.
DR MEROPS; A02_020; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF000075; RNaseH; 1.
DR Pfam; PF000665; rve; 1.
DR Pfam; PF000077; RVP; 1.

DR Pfam: PF00078; RVT_1; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50879; RNASE_H; 1.
SQ SEQUENCE 1145 AA; 128010 MW; 90C1A495E9D95B88 CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1145
Score: 6026.00 Matches: 1143
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 2
Query Match: 97.5% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q9XSN8_PIG (1-1145)

QY 1 ATGGGTGCCACAGGGCAACAACAGTATCCATGGAGCTACCCGGAAGAACAGTTGACTTGGGA 60
DB 1 MetGlyAlaThrGlyGlnGlnTyrProTrpThrThrArgArgThrValAspLeuGly 20

QY 61 GTGGAGCGGTAAACCCACTCGTTCTGTGTACATCTAGTCCAGCAGCCCTCTTAGGT 120
DB 21 ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly 40

QY 121 AGAGACTTATTGACCAAGATCGGAGCACAAAATTTCTTTTGAACAAGGGAACACAGAAAGTG 180
DB 41 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal 60

QY 181 TCTGGAATAACAAACCTATCACTGTGTGTGACCTCCATAGATGAGGAATATCGACTA 240
DB 61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 80

QY 241 TACTCTCCCTAGTAAAGCCGATCAAAATATACAAATTCGTGTGGACAGTTCCTCCCAA 300
DB 81 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln 100

QY 301 GCCTGGGAGAAAACCGAGGATGGTTGGCAAGCAAGTTCCTCCCAACAAGTTATTCAA 360
DB 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120

QY 361 CTGAAGGCCAGTGCACACCAAGTGTCACTAGACAGATACCCCTTGAGTAAAGAAGTCAA 420
DB 121 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln 140

QY 421 GAAGGAATTCGGCGCATGTCCAAAGATTATCAACAGGCGATCCTAGTTCCTGTCCAA 480
DB 141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 160

QY 481 TCTCCTCGAATACTCCCTCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCA 540
DB 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180

QY 541 GTACAGACTTGAGAGGTCAATAAACGGGTGCAGGATATACACCCCAACAGTCCCGAAC 600
DB 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200

QY 601 CCTTATACCTCTGTGTGCTCTCCACCCCAAGCTGATACAGAGCTGGTATACAGTATCGACTTA 660
DB 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu 220

QY 661 AAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAA 720
DB 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240

QY 721 TGGAGAGATCCAGGTACGGGAAGAACCGGGCAGTCTACCTGGACCCGACTGCCCAAGGG 780
DB 241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260

QY 781 TTCAGACTCCCGACCATCTTTGACCAAGCCCTACACAGAGCCTGGCCCACTTCAGG 840
DB 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280

QY 841 ATCCAAACACCTCTCAGGTACACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCC 900

DB 281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAla 300
QY 901 ACCAAACAGGACTCTCTAGAGGCACCAAGCAGCTACTGCTGGAAATGTCTGACCTAGGC 960
DB 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 320

QY 961 TACAGAGCCTCTGCTAAGAAGGCCAGATTTCAGAGAGAGAGGTAAACATCTTGGGGTAC 1020
DB 321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 340

QY 1021 AGTTTGGGACGCGGACGATGGCTGACGAGGACGCGAAGAAACATGTAGTCCAGATA 1080
DB 341 SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 360

QY 1081 CCGSCCCCAACCACAGCCAAACAATGAGAGATTTTGGGACAGCTGGATTTTGAGA 1140
DB 361 ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg 380

QY 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAAACCAAGAAAAA 1200
DB 381 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 400

QY 1201 GGGGAATCTCTGGGCTCTCTGAGCACAGAGGCATTTGATGTATCAAAAAGCCCTG 1260
DB 401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420

QY 1261 CTGAGCCACCTGCTCTGGCCCTCCCTGAGCTAACTAAACCTTTTACCTTTATGTGAT 1320
DB 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440

QY 1321 GAGCGTAAGGAGTAGCCCGGAGATTTTAAACCCAAACCTTAGGACCATGAGAGAGCCT 1380
DB 441 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 460

QY 1381 GTCCCTACTGTCAAAAGAGCTCGATCTGTAGCCAGTGGTGGCCCATATGCTGAAG 1440
DB 461 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTyrProIleCysLeuLys 480

QY 1441 GCTATCCAGCTGTGGCCATCTGCTCAAGAGCCTGACAAATTCATCTTCGGACAGAT 1500
DB 481 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500

QY 1501 ATAACTGTATAGCCCCCATGCTTGGAGAACATCGTTCGGCAGCCCCCAGACCGATGG 1560
DB 501 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 520

QY 1561 ATGACCAAGCCCGCATGACCCATCATCAAGCCTGCTTCTCACAGAGAGGTCACGTTTC 1620
DB 521 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540

QY 1621 GCTCCACGCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGAACCACTG 1680
DB 541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560

QY 1681 ACTCATGATTCCTCACTATTGATTGAGGAGACTTGGGGTCCGCAAGACCTTACAGAC 1740
DB 561 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580

QY 1741 ATACCGCTGACTGAGAGAGTCTAACTGTGTTCTACCGGAAGCAGGATATGTGGTGAA 1800
DB 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600

QY 1801 GGTAAAGGATGGCTGGGGCGGTGTGACGGGACCCGACGATCTGGCCGACGAGC 1860
DB 601 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpPalaSerSer 620

QY 1861 CTGCGGAAGAACTTCAGCAAAAAGGCTCAGTCTATGGCCCTCAGCAAGCTTTGGCG 1920
DB 621 LeuProGlyGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640

QY 1921 CTGCGCGAAGGAAATCCATAAACATTTATACGACAGCAGGTATGCTTTCGCACTGCA 1980
DB 641 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660

QY 1981 CACGTACATGGGGCATCTATAACAAAGGGGTGGCTTACCTCAGCAGGAGGGAATA 2040
 DB 661 HisValHisGlyAlaIlefyLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
 QY 2041 AAGAACAAAGAGGAATTTCTAAGCCTATTAGAAGCCGTACATTTTACCAAAAGGCTAGCT 2100
 DB 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
 QY 2101 ATTATACACTCTCTGGACATCAGAAAGCTAAGATCTCATATCCAGAGAAACAGATG 2160
 DB 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
 QY 2161 GCTGACCGGTTGTCAGACGAGCCAGGGTGTAAACCTTCCTGCTTAAATAGAAATG 2220
 DB 721 AlaAspArgValAlaLysGlnAlaGlnGlyValAsnLeuLeuProIleIleGluMet 740
 QY 2221 CCCAAAGCCCAAGAACCCAGACGACGTACACCCCTAGACGCTGGCAGAGTAAAAAG 2280
 DB 741 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
 QY 2281 ATAGACCACTTCTCTGACACTCCGGAAGGACCTGTATATCTCAGATGGGAAGAAATC 2340
 DB 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
 QY 2341 CTGCCCCACAAAGAGGTTAGAAATATGTCCACAGATACATCGTCTAAACCCACCTAGGA 2400
 DB 781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
 QY 2401 ACTAAACACCTGCAGCAGTGGTGTGAGAACATCCCTTATCATGTTCTGAGGCTACACGGA 2460
 DB 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
 QY 2461 GTGCTCACTCGGTGGTCAACATGTGTGCTCCGTCAGCTGGTAAATGCTTAATCCTTCC 2520
 DB 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
 QY 2521 AGAATGCTCCAGGAAGAGACTAAGGGGAAGCCACCGCGCTCACTGGGAAGTGAC 2580
 DB 841 ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
 QY 2581 TTCACTGAGTAAGCCGGCTAAATACGGAACAAATACCTATTGTTTGTAGACACC 2640
 DB 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
 QY 2641 TTTTCAGATGGGTAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGCTAAA 2700
 DB 881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValAlaLys 900
 QY 2701 AAAATACTGGAAGAAATTTTCCAGATTTTGGAAATCTAAGGTAATAGGCTCAGACAAT 2760
 DB 901 LysIleLeuGluGluIlePheProArgPheGlyLeProLysValIleGlySerAspAsn 920
 QY 2761 GGTCCAGCTTTGTTGCCAGGTAAGTCAAGGAGCTGGCCCAAGATATTTGGGATGATGG 2820
 DB 921 GlyProAlaPheValAlaHisValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
 QY 2821 AAATGCTATTGTGCATACAGACCCCAAGCTCAGGACAGTACAGAGGATGAATAGACC 2880
 DB 941 LysLeuHisCysAlaTyrArgProGlnInsSerGlyGlnValGluArgMetAsnArgThr 960
 QY 2881 ATTAAGAGACCCCTTACTAAATTTGACCGCGAGACTGGCGCTTAATGATTTGATAGCTCTC 2940
 DB 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
 QY 2941 CTGCCCCCTTTGCTTTTAGGGTTAGGAACACCCCTGGACAGTTTGGCTTGACCCCTAT 3000
 DB 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
 QY 3001 GAATTAATCTACGGGGGACCCCGCCATTCGTAGAAATCTCTCTGTACATAGTGTGAC 3060
 DB 1001 GluLeuLeuTyrGlyGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp 1020

QY 3061 GTGCTGCTTTTCCACGCTTTTGTCTTAGGCTCAAGGACTTGTAGTGGGTGAGACAACGA 3120
 DB 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTyrValArgGlnArg 1040
 QY 3121 GCGTGGAGGCAACTCCGGAGGCTTACTCAGGAGGAGACTTGCAGATCCCACTCGT 3180
 DB 1041 AlaTrpArgGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnIleProHisArg 1060
 QY 3181 TTCCAAGTGGAGATTCACTTACGTCTAGTCTAGACGCCACCGTCAGGAAACCTCGAGACTCGG 3240
 DB 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
 QY 3241 TGAAGGCGCTTATCTCTGACTTTTGCACACCAACGCGTGTGAAGTCGAAGGAATC 3300
 DB 1081 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
 QY 3301 TCCACCTGGATCCATGTCATCCAGTTTAAACCGCGCCACTCCCGATTCGGGTGGAAA 3360
 DB 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1120
 QY 3361 GCCGAAAGACTGAAATCCCTTAAGCTTTCGCTCCATCGCTGGTTCCTTACTCTGTC 3420
 DB 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
 QY 3421 AATAACCTCTCAGAC 3435
 DB 1141 AsnAsnLeuSerAsp 1145
 RESULT 2
 Q5Q08_9GAMR PRELIMINARY; PRT; 1193 AA.
 ID Q5Q08_9GAMR PRELIMINARY; PRT; 1193 AA.
 AC Q5Q08;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Pol protein (Fragment).
 OS Porcine endogenous retrovirus C/A.
 OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Gammaretrovirus; 1-Mammalian type C virus group.
 OX NCBI_TaxID=286072;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=A14/220;
 RX PubMed=15564496; DOI=10.1128/JVI.78.24.13880-13890.2004;
 RA Bartosch B., Stefanidis D., Myers R., Weiss R., Patience C.,
 Takeuchi Y.;
 RT "Evidence and consequence of porcine endogenous retrovirus
 recombination.";
 RT J. Virol. 78:13880-13890(2004).
 RL -1- SIMILARITY: Contains 1 reverse transcriptase domain.
 CC EMBL; AV570980; AAT77167.1; -; Genomic DNA.
 DR GO: 0004190; F.aspartic-type endopeptidase activity; IEA.
 DR GO: 0003677; F.DNA binding; IEA.
 DR GO: 0008233; F.peptidase activity; IEA.
 DR GO: 0004523; F.ribonuclease H activity; IEA.
 DR GO: 0003723; F.RNA binding; IEA.
 DR GO: 0003964; F.RNA-directed DNA polymerase activity; IEA.
 DR GO: 0006310; P.DNA recombination; IEA.
 DR GO: 0006508; P.proteolysis and peptidolysis; IEA.
 DR GO: 0006278; P.RNA-dependent DNA replication; IEA.
 DR InterPro: IPR009007; Pept Aspartc_cat.
 DR InterPro: IPR001995; Peptidase_A2_cat.
 DR InterPro: IPR02156; RNaseH.
 DR InterPro: IPR012337; RNaseH_fold.
 DR InterPro: IPR001584; Rve.
 DR InterPro: IPR004477; RVTse.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT_1; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS50994; INTEGRASE; 1.
 DR PROSITE; PS50879; RNASE_H; 1.

DR PROSITE: PS50878; RT POL: 1.
KW Aspartyl protease; P:cease.
FT NON_TER 1
SQ SEQUENCE 1193 AA; 133310 MW; 9F8EC2AD1A66CE10 CRC64;
Alignment Scores:
Pred. No.: 0 Length: 1193
Score: 5992.50 Matches: 1138
Percent Similarity: 99.7% Conservative: 2
Best Local Similarity: 99.5% Mismatches: 3
Query Match: 96.9% Indels: 1
DB: 2 Gaps: 1
US-10-723-552-3_COPY_2307_5741 (1-3435) x Q5Q08_9GAWR (1-1193)
QY 1 ATGGGTGCCACAGGGCAACACAGATATCCATGAGTACCCGGAAGAACAGTTGACTTGGGA 60
DB 50 MetGlyAlaThrGlyGlnGlnGlnTyrSerTrpThrThrArgArgThrValAspLeuGly 69
QY 61 GTGGGACGGTAACCCACTCGTTCTGGTGCATACCTGAGTCCCGACACCCCTCTTAGGT 120
DB 70 ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly 89
QY 121 AGAGACTTATTGACCAAGATGGAGCACAAAATTTCTTTTGAACAGGGAAACCCAGAAGTG 180
DB 90 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal 109
QY 181 TCTGCAATAACAAACCTATCACTGTGTGTGACCTTCCAAATTAGATGACGAATATCGACTA 240
DB 110 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 129
QY 241 TACTCTCCCTAGTAAGCCGTGATCAAAATATACAAATTCGTGTGGACAGTTCCTCCAA 300
DB 130 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln 149
QY 301 GCCTGGGCAGAAACCGCAGGAGTGGTTTGGCAAGCAAGTTTCCCCCAACAAGTTATTCAA 360
DB 150 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 169
QY 361 CTGAAGGCCAGTGCACACAGTGTCACTGACAGACAGTACCCTGTAGTAAAGAAGCTCAA 420
DB 170 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln 189
QY 421 GAAGGAATTCGGCCGATGTCCTCAAGATTAATCCACAGGCGATCCTAGTTCCTGTCCAA 480
DB 190 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 209
QY 481 TCTCCTCGAATACCTCCCTGCTACCGGTAGAAAGCCTGGGACTTAATGACTATCGACCA 540
DB 210 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 229
QY 541 GTACAGGACTTGAGAGGTCAATAACCGGTGCAGGATATACACCCCAACAGTCCCGAAC 600
DB 230 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 249
QY 601 CCTATACCTCTGTGTGCTCTCCACCCCAAGAGCTGGTATACAGTATTGGACTTA 660
DB 250 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu 269
QY 661 AAGGATGCTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTGCCTTCGAA 720
DB 270 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 289
QY 721 TGGAGAGATCCAGGTACCGGAAGAACCGGAGCTCACCTGAGACCCGACTGCCCAAGGG 780
DB 290 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 309
QY 781 TTCAGAACTCCCGACCATCTTTGACCAAGCCCTACACAGACCTGGCCCACTTCAGG 840
DB 310 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 329
QY 841 ATCCAACACCTCTCAGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGGCC 900

DB 330 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAla 349
QY 901 ACCAAACAGGACTGCTTAGAAGGACGACGAGCACTACTGCTGGAATTGCTGACCTAGGC 960
DB 350 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuGluLeuSerAspLeuGly 369
QY 961 TACAGAGCCTCTGCTAAGAGGCCAGATTGTCAGGAGAGAGAGTAAACATCTCTGGGGTAC 1020
DB 370 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 389
QY 1021 AGTTTGGCGGACGGCAGCGATGGCTACCGAGGACGCGAAGAAACCTGTAGTCCAGATA 1080
DB 390 SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 409
QY 1081 CCGGCCCCAACCCACAGCCCAACAAATGAGAGAGTTTGGGACAGCTGGATTTTGCAGA 1140
DB 410 ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg 429
QY 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCCACTCTACCCGCTAACCAAGAAAAA 1200
DB 430 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 449
QY 1201 GGGGAATTCCTCTGGGCTCTGACGACAGAGGATTTGATGCTATCAAAAAAGCCCTG 1260
DB 450 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 469
QY 1261 CTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTTAAACCTTTACCTTTATGCTGAT 1320
DB 470 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 489
QY 1321 GAGCGTAGGAGTAGCCGGGGAGTTTAAACCCCAACCCCTAGGACCATGAGAGAGACCT 1380
DB 490 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 509
QY 1381 GTCCCTACCTGTCAAAGAAGCTCGATCCTGTAGCCAGTGGTTGGCCCATATGCTGAAG 1440
DB 510 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTyrProIleCysLeuLys 529
QY 1441 GCTATCGCAGTGTGGCCATCTGTCAGGACCTGACAAATTCGACTTTGGGACAGAT 1500
DB 530 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 549
QY 1501 ATAAGTGTATAGCCCCCATGCTGAGAAACATCGTTCGGCAGCCGCCAGACCGATGG 1560
DB 550 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 569
QY 1561 ATGACCAACGCCCATGACCCACTATCAAAGCCTGCTTTCACAGAGAGGTCACGTTTC 1620
DB 570 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 589
QY 1621 GCTCCACGACCGCTCTCAACCCCTGCCACTCTTCTGCTGAGAGACTGATGAACCCAGTG 1680
DB 590 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 609
QY 1681 ACTCATGATTCGCATCAACTATTGATGAGAGACTTGGGGTCCGCAAGGACCTTACAGAC 1740
DB 610 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 629
QY 1741 ATACCGCTGACTGGAAGATGCTAACTCTGGTTCACTGACGGAAGCAGCTATGTGGTGAA 1800
DB 630 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 649
QY 1801 GGTAAGGATGCTGGGGCGGGTGTGTGACGGGACCCGACGACTGCGCCGAGCAGC 1860
DB 650 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 669
QY 1861 CTGCGGGAAGAACTTCAGCAAAAAGCTCAGCTCATGCGCCCTCAGCAAGCTTTTGGCG 1920
DB 670 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 689
QY 1921 CTGCGCGAAGGAAATCCATAAACATTATACGACAGCAGGTATGCTTTGCGACTGCA 1980
DB 690 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 709

QY	1981	CACGTACATGGGGGCATCTATAAACAAGGGGTGGCTTACCTCAGCAGGAGGAAATA	2040
Db	710	HisValHisGlyAlaIleTyLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle	729
QY	2041	AAGACAAAGAGGAATCTTAAGCTATTAGACCGGTACATTACCAAAAAGCTAGCT	2100
Db	730	LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla	749
QY	2101	ATTATACACTCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATG	2160
Db	750	IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet	769
QY	2161	GCTGACCGGGTTGCAAGCAGGACCCAGGGGTGTTAACTTCTGCCCTTATATAGAAATG	2220
Db	770	AlaAspArgValAlaLysGlnAlaLagInGlyValAsnLeuLeuProIleGluMet	789
QY	2221	CCCAAGCCCGACAGCAGCAGCTACACCTTAGNAGACTGGCAAGAGATAAAAAG	2280
Db	790	ProLysAlaProGluProArgArgGlnTyThrLeuGluAspTrpGlnGluIleLysLys	809
QY	2281	ATAGACAGATGTTCTCTGAGACTCCGGAAGGAGCTGCTATACCTCAGATGGGAAGAAATC	2340
Db	810	IleAspGlnPheSerGluThrProGluGlyThrCystyrThrSerAspGlyLysGluIle	829
QY	2341	CTGCCCCCAAAAGAGGTTAGAATATGTCCAAACAGATACATCGCTAAACCCACCTAGGA	2400
Db	830	LeuProHisLysGluGlyLeuGluTyValGlnGlnIleHisArgLeuThrHisLeuGly	849
QY	2401	ACTAAACACCTGCAGCAGCTGGTCAGAACATCCCTTATCATGTCTGAGGCTACAGGA	2460
Db	850	ThrLysHisLeuGlnGlnLeuValArgThrSerProTyThrHisValLeuArgLeuProGly	869
QY	2461	GTGCTGACTCGGTGGTCAAACTTGTCGCCCTGCAGCTGGTTAATGCTTAATCCTTCC	2520
Db	870	ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer	889
QY	2521	AGAATGCTCCAGGAAGAGACTAAGGGGAAGCCACCGAGCGCTCACTGGGAAGTGAC	2580
Db	890	ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp	909
QY	2581	TTCACTCAGGTAAAGCCGGCTAAATACGGAACAATAACCTATTGGTTTGTAGACACC	2640
Db	910	PheThrGluValLysProAlaLysTyGlyAsnLysTyThrLeuLeuValPheValAspThr	929
QY	2641	TTTTCAAGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACTTCAACCGGTGGCTAAA	2700
Db	930	PheSerGlyTrpValGluAlaTyProThrLysLysGluThrSerThrValValAlaLys	949
QY	2701	AAAATACTGGAAGAAATTTTCCAAGATTTGGAAATACCTAAGGTAAATAGGGTCAGACAAT	2760
Db	950	LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn	969
QY	2761	GGTCCAGCTTTGTGCGCAGGTAAAGTACGAGGCTGGCCAGATATATGGGGATGATGG	2820
Db	970	GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp	989
QY	2821	AAACTGCATGTGCATACAGACCCCAAGCTCAGACAGGTAGAGAGGATGAATAGAAC	2880
Db	990	LysLeuHisCysAlaTyArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr	1009
QY	2881	ATTAAGAGACCCCTTACTAAATTGACCGCGAGACTGGCGGTAAATGATGGATAGCTCTC	2940
Db	1010	IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu	1029
QY	2941	CTGCCCTTTGTGCTTTTATAGGTTAGAAACACCCCTGAGCAGTTTGGCTGACCCCTTAT	3000
Db	1030	LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTy	1049
QY	3001	GAATTACTCTACGGGGACCCCCCATTTGGTAGAAATTCCTCTGTACATAGTCTGAC	3060
Db	1050	GluLeuLeuTyrgLysGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp	1069
QY	3061	GTGCTGCTTTTCCAGCCTTTTCTTAGGCTCAAGGCATTTAGTGGGTGAGACAACGA	3120
Db	1070	ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg	1089
QY	3121	GCCTGGAGGCAACTCCGGGAGGCTACTCAGGAGGAGAGACTTGCAGATCCCATCGT	3180
Db	1090	AlaTrpLysGlnLeuArgGluAlaTySer---GlyGlyAspLeuGlnValProHisArg	1108
QY	3181	TTCCAAGTGGGAGATTCACTTACGTCAGCGCCACCGTGCAGGAAACCTCGAGACTCGG	3240
Db	1109	PheGlnValGlyAspSerValTyValArgArgHisArgAlaGlyAsnLeuGluThrArg	1128
QY	3241	TGGAAGCGCCCTTATCTCTGCTACTTTTCACCAACCAACCGCTGTGAAGTCAAGGAATC	3300
Db	1129	TrpLysGlyProTyThrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle	1148
QY	3301	TCCACCTGGATCCATGTCATCCACGTTAAACCGCGCCACCTCCCGATTTCGGGTGAAA	3360
Db	1149	ProThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys	1168
QY	3361	GCCGAAAGACTGAAAATCCCTTAAGCTTCGCTCCATCGCGTGTTCCTTACTCTGTC	3420
Db	1169	AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTySerVal	1188
QY	3421	AATAACCTCTCA 3432	
Db	1189	AsnAsnSerSer 1192	
RESULT 3			
Q4VFZ2_9GAMR			
ID	Q4VFZ2_9GAMR PRELIMINARY;	PRT;	1145 AA.
AC	Q4VFZ2;		
DT	13-SEP-2005 (TrEMBLrel. 31, Created)		
DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)		
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)		
DE	Pol protein.		
GN	Namespol;		
OS	Porcine endogenous retrovirus C/A.		
OC	Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;		
OC	Gammaretrovirus; 1-Mammalian type C virus group.		
NCBI_TaxID=286072;			
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=C/A 4;		
RA	Denner J., Karlas A., Votteler Jr.;		
RT	"Nucleotide sequence and extended characterization of a high passage		
RL	human cell-adapted recombinant FERV-C/A.";		
DR	Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.		
EMBL; AY953542; AAY28928.1; -; Genomic DNA.			
SQ	SEQUENCE 1145 AA; 127925 MW; 62014B9629B04051 CRC64;		
Alignment Scores:			
Pred. No.:	0	Length:	1145
Score:	5986.00	Matches:	1134
Percent Similarity:	99.6%	Conservative:	5
Best Local Similarity:	99.1%	Mismatches:	5
Query Match:	96.8%	Indels:	0
DB:	2	Gaps:	0
US-10-723-552-3_COPY_2307_5741 (1-3435) x Q4VFZ2_9GAMR (1-1145)			
QY	1	ATGGGTGCCACAGCGGCAACAACAGTAGTATCCATGAGCTACCCGAAAGACAGTTGACTTGGGA	60
Db	1	MetGlyAlaThrGlyGlnGlnGlnTyProTrpThrThrArgThrValAspLeuGly	20
QY	61	GTGGGACGGGTAAACCACTGCTTTTCTGGTCATACCTGAGTCCGACACCCCTCTTAGGT	120
Db	21	ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly	40
QY	121	AGAGACTTATTGACCAAGATGGGAGGACAAATTTCTTTTGAACAAGGAACACAGATG	180
Db	41	ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal	60

181 TCTGCAATAACAAACCTATCAGTGTGTGGTACCCTCCAAATTAGATGACGAATATCGACTA 240
61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 80
241 TACTCTCCCTAGTAAGCCTGATCAAAATATACAAATTCTGGTTGGAAACAGTTCCCCCAA 300
81 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln 100
301 GCCTGGGAGAAACCGCAGGAGTGGTTGGCAAGCAAGTTCCCCCAAGATTATTCAA 360
101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120
361 CTGAGGCCAGTGGCACACAGTGTCACTCAGACAGTACCCCTTGAGTAAAGAGCTCAA 420
121 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln 140
421 GAAGGAATTCGGCCGATGTCAAAGATTAAATCAACAGGCGATCCTAGTTCTCTGTCCAA 480
141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlyIleLeuValProValGln 160
481 TCTCCTCGGAATACCTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCA 540
161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
541 GTACAGGACTTGAGAGGTCAATAAACGGGTGCAGGATATACACCCCAACAGTCCCGAAC 600
181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200
601 CTTTATAACCTCTTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGACTTA 660
201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu 220
661 AAGGAGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAA 720
221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
721 TGGAGAGATCCAGGTACCGGAAGAACCGGCGAGCTCACCTGACCCGACTGCCCAAGGG 780
241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
781 TTCAAGAACTCCCGACCATCTTTGACGAAGCCCTTACACAGAGACTTGGCCAACTTCAGG 840
261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280
841 ATCCAAACCTCTCAGGTGACCTCTCCAGTACGTGGATGACCTCTCTGGCGGAGCC 900
281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 300
901 ACCAAACAGGACTGCTTGAAGGCACGAGGCACCTACTGCTGGAATTCCTGACCTAGGC 960
301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 320
961 TACAGAGCTCTGCTTGAAGGCCCAAGATTTGCAGGAGAGAGTTAACATACTTGGGGTAC 1020
321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgAlaGluValThrTyrLeuGlyTyr 340
1021 AGTTTGGGAGCGGCGAGCGATGCTGACGAGGACCGAGCAAACTGTAGTCCAGATA 1080
341 SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 360
1081 CCGGCCCCCAACACAGCCAAACAAATGAGAGATTTTTGGGACAGCTGGATTTTTCAGA 1140
361 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 380
1141 CTGTGGATCCCGGGTTTGGCACTTAGCAGCCCCCACTCTACCCGCTAACCAAGAAAAA 1200
381 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 400
1201 GGGGAATTCCTCTGGGCTCTGAGCACGAGAGGCAATTTGATGCTATCAAAAAGCCCTG 1260
401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
1261 CTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCCCTTTATCGGAT 1320

421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440
1321 GAGGTAAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGACACCATGGAGAGACCT 1380
441 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgPro 460
1381 GTCCCTACCTCTCAAGAGCTCGATCCTGTAGCTGTGGTGGCCCATATGCTCTGAAG 1440
461 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys 480
1441 GCTATCCAGCTGTGGCCATACCTGCTCAAGACCTCTGACAAATTGACTTTGGGACAGAA 1500
481 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500
1501 ATAACTGTAATAGCCCCCATATGCTGGAGAAATCCTGTGGCAGCCGCCAGACCGATGG 1560
501 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 520
1561 ATGACCAACCGCCGATGACCCACTATCAAGCCTCTTCTGCTGGAAGACTGTAAACCACTG 1620
521 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540
1621 GCTCCACGACCGCTCTCAACCCCTGCCACTCTTCTGCTGGAAGACTGTAAACCACTG 1680
541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGlnThrAspGluProVal 560
1681 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGTCCGCAAGGACCTTACAGAC 1740
561 ThrHisaspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580
1741 ATACCGCTGACTGGAGAAGTGTAACTGTGTTCACTGACGGAACAGCTATGTGTGGAA 1800
581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
1801 GGTAAAGGATGGTGGCGCGGTGTGTGACCGGACCCCGCACGATCTGGCCGACGACG 1860
601 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
1861 CTGCGGGAAGAACTTCAGCAAAAAGCTGAGCTCATGGCCCTCAGCCAAAGCTTTGGCG 1920
621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
1921 CTGCGCGAAGGAATCCATAAACATTTATACGACACGACGAGTATGCTTTGCCACTGCA 1980
641 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
1981 CACGTACATGGGCGCATCTATAAAACAAAGGGGTGTGCTTACCTCAGCAGGAGGGAATA 2040
661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
2041 AAGAACAAGAGGAAATCTTAAGCCTATTAGAAGCTGATACATTTACCAAAAAGCTAGCT 2100
681 LysAsnLysGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 700
2101 ATTATACACTGCTCTGGACATCAGAAAGCTCTCATATCCAGAGGAACACAGATG 2160
701 IleIleHisCysProGlyHisGlnLysAlaLysAspProIleSerArgGlyAsnGlnMet 720
2161 GCTCAGCCGGTGTGCAAGCAGCGACCCAGGGTGTAACTTCTGCTCTATATAGAAATG 2220
721 AlaAspArgValAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProMetIleGluThr 740
2221 CCCAAAGCCCGAAGCCAGACGACGTACACCTTAGAGACTGGCAGAGATAAAAAAG 2280
741 ProlysAlaProGluProGlyArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
2281 ATAGACCAAGTTCTCTGAGACTCCGGAAGGACCTGTATATCTCAGATGGGAAGGAAATC 2340
761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
2341 CTGCGCCCAAGGAAGGGTTAGAAATATGTCCAACAGATACATCGTCTTAACCCCACTAGA 2400

Db 781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
 QY 2401 ACTAAACACCTGCAGCAGTGTGTGAGAACATCCCTTATCATGTTCTCAGGCTACACGGA 2460
 Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
 QY 2461 GTGGCTGACTGGTGGTCAAACTGTGTGCTGCCAGCTGGTTAAATGCTTAATCTTCC 2520
 Db 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
 QY 2521 AGAATGCTCTCAGGGAAGAGACTAAGGGCAAGACCCAGCGCTCCTGCGAAGTGAC 2580
 Db 841 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpValValAsp 860
 QY 2581 TTCCTAGGTAAGCCCGCTAAATACGGAACAATACTATTGTTTGTAGACACC 2640
 Db 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
 QY 2641 TTTTCAGGATGGTAGAGCTTATCTTACTAAGAAGAGACTTCAACCGTGGTCTTAA 2700
 Db 881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
 QY 2701 AAAATACCTGGGAAGAAATTTTCCAGATTTTCCAAATACCTAAGGTAAATAGGTCAGACAT 2760
 Db 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
 QY 2761 GGTCCAGCTTTTGTGTCAGTAACTAGTCAGGAGCTGGCCAAAGATATTGGGATTTGATGG 2820
 Db 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
 QY 2821 AAATGCTGATTTGCTGATCAGACCCCAAGCTCAGACAGGTAGAGAGGATGAATAGAAC 2880
 Db 941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
 QY 2881 ATTTAAAGACCTTACTTAATTTGACCGCGAGACTGGCGTTAATGATGATAGCTCTC 2940
 Db 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAlaAsnAspTrpIleAlaLeu 980
 QY 2941 CTGCGCTTTTGTGCTTTTGTAGGGTTAGGAACACCCCTGACAGTATTGGGCTGACCCCTAT 3000
 Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
 QY 3001 GAATTAATCTAGGGGGAGACCCCTTCTAGGCTCAAGGCTTGTAGAAATGCTTCTGTACATAGTCTGAC 3060
 Db 1001 GluLeuLeuTyrGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
 QY 3061 GTGCTGCTTTCCAGCTTTTCTTCTAGGCTCAAGGCTTGTAGTGGTGGTGGAGACAGA 3120
 Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
 QY 3121 CGCTGGAGGCAACTCCGGGGAGCTTCTCAGGAGGAGAGACTTGCAGATCCACATCGT 3180
 Db 1041 AlaTrpArgGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnIleProHisArg 1060
 QY 3181 TTCCAAGTGGAGATTGACTTACGTTAGACCCAGCGCTGAGGAAACCTCGAGACTCGG 3240
 Db 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
 QY 3241 TGGAGGCGCTTATCTCGTACTTTTGCACACCAAGCGCTGTGAAGTCAAGGATC 3300
 Db 1081 TrpLysGlyProTyrHisValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
 QY 3301 TCCACCTGGATCCATGCTCCAGCTTAAACCGCGCCACCTCCCGATTTCGGGTGGAAA 3360
 Db 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1120
 QY 3361 GCCGAAAGACTGAAATCCCTTAAAGTTCGCTCCATCGCTGGTTCCTTACTCTGTC 3420
 Db 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
 QY 3421 AATAACCTCTCA 3432
 Db 1141 AsnAsnPheSer 1144

RESULT 4

Q90RL9_9GAMR
 ID Q90RL9_9GAMR PRELIMINARY; PRT; 1193 AA.
 AC Q90RL9_2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polymerase (Fragment).
 GN Name=pol;
 OS Porcine endogenous type C retrovirus.
 OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 CC Gammaretrovirus.
 OX NCBI_TaxID=70540;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 EX MEDLINE=21256017; PubMed=11356953;
 EX DOI=10.1128/JVI.75.12.5465-5472.2001;
 RA Krach U., Fischer N., Czauderna F., Toenjes R.R.;
 RT "Comparison of replication-competent molecular clones of porcine
 RT endogenous retrovirus class A and class B derived from pig and human
 cells.";
 RL J. Virol. 75:5465-5472(2001).
 CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
 DR EMBL; AJ293656; CAC39617.1; -, Genomic_DNA.
 DR HSSP; P03355; 1I6J.
 DR MEROPS; A02.020; -.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR009007; Pept_Aspart_cat.
 DR InterPro; IPR001595; Peptidase_A2_cat.
 DR InterPro; IPR001256; RNaseH.
 DR InterPro; IPR012337; RNaseH_fold.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS50994; INTEGRASE; 1.
 DR PROSITE; PS50879; RNASE_H; 1.
 DR PROSITE; PS50878; RT_POL; 1.
 KW Aspartyl protease; protease.
 FT NON_TER 1
 SQ SEQUENCE 1193 AA; 133225 MW; 8FC1F419CB600332 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 1193
 Score: 5914.50 Matches: 1119
 Percent Similarity: 98.9% Conservative: 12
 Best Local Similarity: 97.8% Mismatches: 12
 Query Match: 95.7% Indels: 1
 DB: 2 Gaps: 1

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q90RL9_9GAMR (1-1193)

QY 1 ATGGGTGCCACAGGGCAACAGTATCCATGACTACCCGAGAACAGTTGACTTGGGA 60
 Db 50 MetGlyAlaThrGlyGlnGlnTyrProThrThrArgThrValAspLeuGly 69
 QY 61 GTGGGCGGTAAACCACTCGTTTCTGGTCATACCTGAGTCCAGCACCCCTCTTAGGT 120
 Db 70 ValGlyArgValThrHisPheLeuValIleProGluCysProAlaProLeuLeuGly 89
 QY 121 AGAGACTTATTGACCAAGATGGGAGCAAAATTTCTTTTGAACAGGGAACCAAGATG 180

Db 90 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal 109
Qy 181 TCTGCAATAAACAACCTATCACTGTGTGTGACCTCCAAATTTAGATGACGAATATCGACTA 240
Db 110 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 129
Qy 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATTCATCTGGTGGACAGTTTCCCCAA 300
Db 130 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln 149
Qy 301 GCCTGGGAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCCAAGATTATTCAA 360
Db 150 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 169
Qy 361 CTGAAGGCCAGTGCCACACCAAGTGTCACTGACAGACAGTACCCCTTGAGTAAAGAGCTCAA 420
Db 170 LeuLysAlaSerAlaAlaProValSerValArgGlnTyrProLeuSerLysGluAlaArg 189
Qy 421 GAAGGAATTCGGCCGATGTCAAAGATTATCCACAGGGCATCTAGTTCTCTGTCCAA 480
Db 190 GluGlyIleLeuArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 209
Qy 481 TCTCCTCGAATACTCCCTGCTACCGGTAGAAAGCCTGGGACTTAATGACTATCGACCA 540
Db 210 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 229
Qy 541 GTACAGGACTTCGAGAGGTCAATAAACGGGTGCAGGATATACACCCCAAGTCCCGAAC 600
Db 230 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 249
Qy 601 CTTATACCTCTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGACTTA 660
Db 250 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu 269
Qy 661 AAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAA 720
Db 270 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 289
Qy 721 TGGAGAGATCCAGGTPACGGGAAGAACCGGGCAGCTCACCTGGACCGCATGCCCAAGGG 780
Db 290 TrpArgAspProGlyAlaGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 309
Qy 781 TTCAGAACTCCCGACCATCTTTGACCAAGCCCTACACAGACCTGGCCCACTTCAGG 840
Db 310 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 329
Qy 841 ATCCAAACACCTCAGTGCACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGAGCC 900
Db 330 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 349
Qy 901 ACCAAACAGGACTGCTTAGAGGCAAGGACACTACTGCTGGAATTCCTGACCTAGGC 960
Db 350 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuLeuLeuLeuLeuLeu 369
Qy 961 TACAGAGCTCTGCTAAGAGCCCAAGATTTGCAGAGAGGATTAACATCTGGGTAC 1020
Db 370 TyrArgAlaSerAlaLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 389
Qy 1021 AGTTTGGGGAGGGCAGCGATGGCTGACGGAGGACCGAAGAACTGTAGTCCAGATA 1080
Db 390 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysArgThrValValGlnIle 409
Qy 1081 CCGGCCCAACACACGCAACAAATGAGAGAGTGTTCGGGACAGCTGGATTTGCGAGA 1140
Db 410 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 429
Qy 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAA 1200
Db 430 LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGlnLys 449
Qy 1201 GGGGAATTCCTCGGCTCCTGAGCACGAGGATTTGATGCTTATCAAAAAGGCCCTG 1260

Db 450 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 469
Qy 1261 CTGAGCCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTACCTTTATGTGGAT 1320
Db 470 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 489
Qy 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGACCATGAGAGACCT 1380
Db 490 GluArgLysGlyValAlaIleArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 509
Qy 1381 GTCCCTACCTGTCAAAAGAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGCTGAAG 1440
Db 510 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTyrProIleCysLeuLys 529
Qy 1441 GCTATCCAGCTGTGGCCATCTGGTCAAGGACCTGACAAATTCAGCTTTGGGACAGAT 1500
Db 530 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 549
Qy 1501 ATAACTGTATAGCCCCCATGCTTGGAGAACATCGTTGGCAGCCGCCAGACCGATGG 1560
Db 550 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 569
Qy 1561 ATGACCAACGCCCATGACCCACTATCAAAGCTGCTTCTCACAGAGAGGTCAACGTC 1620
Db 570 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 589
Qy 1621 GCTCACAGCGCTCTCAACCTGCCACTCTTCTGCTGAGAGACTGATGACACCGTG 1680
Db 590 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 609
Qy 1681 ACTCATGATGCCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 1740
Db 610 ThrHisaspCysHisGlnLeuLeuIleGluThrGlyValArgLysAspLeuThrAsp 629
Qy 1741 ATACCGCTGACTGAGAAGTGCTAACCTGGTTCTACTGACGGAACAGCTATGTGGTGAA 1800
Db 630 IleProLeuThrGlyGluMetLeuThrTrpPheThrAspGlySerSerTyrMetValGlu 649
Qy 1801 GGTAAAGAGATGGCTGGGGCGGGTGGTGGACGGGACCCGACGATCTGGCCGACGAC 1860
Db 650 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 669
Qy 1861 CTGCGCGAAGGAACCTTCAGCACAAAGGCTGAGCTCATGGCCCTCAGCAAGCTTTGCGG 1920
Db 670 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 689
Qy 1921 CTGCGCGAAGGAATCCATAAACATTATATACGACGACAGGTATGCTTTGCGACTGCA 1980
Db 690 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 709
Qy 1981 CACGTACATGGGCGCATCTATAACAAAGGGGTGGTTCCTCAGCAGGAGGGGAATA 2040
Db 710 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 729
Qy 2041 AAGAAACAAAGAGGAATTTCTAAGCCTATTAGAAGCCGTACATTTTACCAAAAAGGCTAGCT 2100
Db 730 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 749
Qy 2101 ATTATACACTGCTCTGGACATCAGAAAAGCTTCAATATCTCATATCCAGAGGAAACAGATG 2160
Db 750 IleIleHisCysProGlyHisGlnLysAlaLysAspProIleSerArgGlyAsnGlnMet 769
Qy 2161 GCTGACCGGGTTGCCAAGCAGGCGCCAGGGTGTAACTTTCCTCTATATATAGAAATG 2220
Db 770 AlaAspArgValAlaLysGlnAlaGlnGlyValAsnLeuLeuProMetIleGluThr 789
Qy 2221 CCCAAAGCCCGAAGCCAGACAGTACACCTTAGAAGACTGCGAAGATGATAAAAAG 2280
Db 790 ProLysAlaProGluProGlyArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 809
Qy 2281 ATAGACCAAGTTCTCTGAGACTCCGAAAGGACCTGCTATATACCTCAGATGGGAAGAAATC 2340
Db 810 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 829

QY 2101 ATTATACACTGCTCTGGACATCAGAAAGCTTAAGATCTCATATCCAGAGGAAACAGATG 2160
Db 1274 IleIleHisCysProGlyHisGlnLysAlaLysAsePLeuIleSerArgGlyAseGlnMet 1293
QY 2161 GCTGACCGGGTTGCCAAGCAGCAGCCAGGGGTGTTAACTTCTGCTTAAATGAATG 2220
Db 1294 AlaAsePValAlaLysGlnAlaGlnAlaValAseLeuLeuProIleIleGluThr 1313
QY 2221 CCCAAAGCCCAAGACCCAGACGACAGTACACCTTAGAAGACTGGCAAGAGATAAAAAAG 2280
Db 1314 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAsePTrpGlnGluIleLysLys 1333
QY 2281 ATAGACCAAGTCTCTGACACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATC 2340
Db 1334 IleAsePLeuPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 1353
QY 2341 CTGCCCAACAAGAGGGTGTAGAATATCTCCACAGATACATCTCTAAACCCACTAGGA 2400
Db 1354 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 1373
QY 2401 ACTAAACACCTGCAGCAGTCTGGTCAGAACATCCCCCTATCATGTTCTGAGGCTACCAAG 2460
Db 1374 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 1393
QY 2461 GTGGCTGACTCGTGGTCAAAATGTTGTGCTGCCAGCTGGTTAATGCTAAATCCTTCC 2520
Db 1394 ValAlaAsePLeuValLysHisCysValProCysGlnLeuValAseAlaAseProSer 1413
QY 2521 AGAATGCTCCAGGAGAGACTAAGGGAGACCCAGCGCTCACTGGGAAGTGAC 2580
Db 1414 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAseP 1433
QY 2581 TTCACTGAGTAAAGCCGGCTAAATATACGAAACAAATACCTATTGTTTGTAGACACC 2640
Db 1434 PheThrGluValLysProAlaLysTyrGlyAseLysTyrLeuLeuValPheValAsePThr 1453
QY 2641 TTTTCAGATGGGTAGAGCTTATCTCTAAGAAAGAGACTCAACCGTGGTGGCTAAA 2700
Db 1454 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 1473
QY 2701 AAAATACTGGAGAAATTTTCCAGATTGGAATACCTAAGTAAATAGGTCAGACAT 2760
Db 1474 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAsePAsn 1493
QY 2761 GGTCCAGCTTTGTTGCCAGTAAAGTCAGGACTGGCCAAAGATATTGGGGATTGATTGG 2820
Db 1494 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAsePTrp 1513
QY 2821 AAATGCTATTGTGCATACAGACCCCAAGCTCAGACAGGTAGAGAGATGAATAGAAC 2880
Db 1514 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluAsePMetAsePThr 1533
QY 2881 ATTAAGAGACCTTACTTAATTCACCGGAGACTGGCGTTAATGATTGATAGTCTC 2940
Db 1534 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsePTrpMetAlaLeu 1553
QY 2941 CTGCCCTTTGTGCTTTTAGGGTTAGGAACACCCCTCGACAGTTGGTGGCTGACCCCTAT 3000
Db 1554 LeuProPheValLeuPheArgValArgAsePTrpGlyGlnPheGlyLeuThrProTyr 1573
QY 3001 GAATTACTACTACGGGGACCCCCCAATTTGGTAGAAATTCCTCTGTACATAGTGTGAC 3060
Db 1574 GluLeuLeuTyrGlyGlyProProProLeuValGluIleAsePValHisSerAlaAseP 1593
QY 3061 GTGCTGCTTCCAGACCTTTGTTCTTAGGCTCAAGGCACTTGAAGTGGGTGAGACAGCA 3120
Db 1594 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1613
QY 3121 GCGTGGAGGCAACTCCGGAGGCTACTCAGGAGGAGAGACTTGCAGATCCCACTCGT 3180
Db 1614 AlaTrpLysGlnLeuArgGluAlaTyrSerGlyGluGlyAsePLeuGlnValProHisArg 1633
QY 3181 TTCAAGTGGAGATTCACTAGTCTACGTTAGACGCCCGCTGACGAAACCTCGAGACTCGG 3240

Db 1634 PheGlnValGlyAsePLeuValArgArgHisArgAlaGlyAsePLeuGluThrArg 1653
QY 3241 TGGAGGGCCCTTATCTCGTACTTTTGACACACCAACCGCTGTGAAGTGAAGGAATC 3300
Db 1654 TrpLysGlyProTyrLeuValLeuThrThrProThrAlaValLysValGluArgIle 1673
QY 3301 TCCACCTGGATCCATGCCATCCAGCTTAAACCGCGCCACCTCCCGATTCCGGGTGAAA 3360
Db 1674 ProThrTrpIleHisAlaSerHisValLysProAlaProProAsePLeuTrpLys 1693
QY 3361 GCCGAAAGACTGAAATCCCTTAAAGCTTCGCTCCATCCGCTGGTTCCTTACTCTC 3420
Db 1694 AlaGluLysThrGluAsePLeuProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1713
QY 3421 AATAACCTCTCA 3432
Db 1714 AsnAsePLeuSer 1717
RESULT 6
Q8J4V6 9GAMR
ID Q8J4V6 9GAMR PRELIMINARY; PRT; 1720 AA.
AC Q8J4V6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DS Gag-pol polyprotein.
OS Porcine endogenous retrovirus B.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=194959;
RN [1]
RP NUCLEOTIDE SEQUENCE
MEDLINE=22172947; PubMed=12185278;
RA Bartosch B., Weiss R.A., Takeuchi Y.;
RT "PCR-based cloning and immunocytological titration of infectious
porcine endogenous retrovirus subgroup A and B";
J. Gen. Virol. 83:2231-2240(2002).
RL EMBL; AY09324; AAM29194.1; -; Genomic_DNA.
DR HSP; P03355; I16J.
DR SMR; Q8J4V6; 7-98.
DR MEROPS; A02.020; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G retro matrix.
DR InterPro; IPR001995; Peptidase A2_cat.
DR InterPro; IPR001969; Pept_Asp_AS_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA_1.
DR Pfam; PF02093; Gag_p30_1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00879; RNASE_H; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.

KW Polyprotein.
SQ SEQUENCE 1720 AA; 193393 MW; 7F9DCC2F5B834D8A CRC64;

Alignment Scores:

Pred. No.: 0 Length: 1720
Score: 5872.00 Matches: 1111
Percent Similarity: 98.7% Conservative: 16
Best Local Similarity: 97.3% Mismatches: 15
Query Match: 95.0% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q8J4V6_9GAMR (1-1720)

```
QY 1 ATGGGTGCACAGGGCAACACGATCATCATGGACTACCCGAAGACGATTCGACTTGGGA 60
DB MetGlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValAspLeuGly 594
QY 61 GTGGACGGGTAAACCACTCGTTCTGCTGATACCTGAGTGCACAGACCCCTCTTAGGT 120
DB ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly 614
QY 121 AGAGACTTATTGACCAAGATGGAGACAAATTTCTTTTGAAACAAGGAAACCAAGATG 180
DB ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 634
QY 181 TCTGCAATAACAACCTTACTCTGCTGTGTGACCTCCCAATTAGATGCAATATCGACTA 240
DB SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 654
QY 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAATTCGTGTTGGAAACAGTTTCCCCAA 300
DB TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln 674
QY 301 GCCTGGGCAGAAACCGCAGGAGTGGTTTGGCAAAGCAAGTTCCTCCCAACAAGTTATTCAA 360
DB AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 694
QY 361 CTGAAGGCCAGTGCACACGAGTGTCACTGACAGACGATCCCTTGAGTAAAGAGCTCAA 420
DB LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg 714
QY 421 GAAGGAATTCGGCCCATGTCCAAAGATTAAATCAACAGGCGATCCTAGTTCTCTGTCCAA 480
DB GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 734
QY 481 TCTCCCTGGAATATCTCCCTGCTACCGGTTAGAAAGCCTGGACCTAATGACTATCGACCA 540
DB SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 754
QY 541 GTACAGGACTTGAGAGGTCATATAACGGGTGCAGGATATACACCCACAGTCCCGAAC 600
DB ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 774
QY 601 CCTTATAACCTCTTGTGTGCTCTCCACCCCAACGGAGCTGATATACAGTATTGGACTTA 660
DB ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeu 794
QY 661 AAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAA 720
DB LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 814
QY 721 TGGAGAGATCAGGTACGGGAAGAACCAGGCGAGCTCACCTGACCCGACTGCCCAAGGG 780
DB TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 834
QY 781 TTCGAAGACTCCCGACCATCTTTGACGAAGCCCTTACACAGAGACTGCCCAACTTCAGG 840
DB PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 854
QY 841 ATCCAAACACCTCAGGTACCCCTCTCTCAGTACGTGGATGACCTGCTTCTCGCGGAGCC 900
DB ILeGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAla 874
```

```
QY 901 ACCAAACAGGACTGCTTTAGAAAGGCACGAAGGCACACTACTGCTGGAATTGCTGACCTAGGC 960
DB ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 894
QY 961 TACAGAGCCTCTGCTAAGAAAGCCAGATTTCGAGGAGAGAGGTAAACATATCTTGGGTAC 1020
DB TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 914
QY 1021 AGTTTGGGGACGGCGAGCGATGGCTACGGAGCAGCGAAGAAACCTGTAGTCCAGATA 1080
DB SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 934
QY 1081 CCGGCCCAACACAGACCAACAATGAGAGAGTTTTCGGGACAGCTGGGATTTTGACA 1140
DB ProIleProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 954
QY 1141 CTGTGGATCCCGGGTTTTCGACCTTAGCAGCCCACTCTTACCCGCTTAACCAAGAAAAA 1200
DB LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 974
QY 1201 GGGGAATTCCTCGGCTCCTGAGCAGCAGAGGACATTTGATGCTATCAAAAAGCCCTG 1260
DB GlyGlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 994
QY 1261 CTGAGCGCACCTGCTCGGCCCTCCCTGACGTAACCTTAAACCTTTTACCTTTATGTGGAT 1320
DB LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 1014
QY 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCAAAACCTTAGCACCATGAGAGAGACCT 1380
DB GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 1034
QY 1381 GTCGCTACCTGTCAAAGAAGCTCGATCCTGACGAGTGGTTGGCCCATATGCTGAG 1440
DB ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys 1054
QY 1441 GCTATCGCAGCTGGCGCATACTGCTCAAGGACCTGACAAATTTGACTTTGGGACAGAA 1500
DB AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 1074
QY 1501 ATAACTGTAATAGCCCCCATGCTTGGAGAACATCGTTTCGGCAGCCCCCAGACCGATGG 1560
DB IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 1094
QY 1561 ATGACCAACGCCCGCATGACCCACTATCAAAAGCTGTTCTTCACAGAGAGGTCACGTT 1620
DB MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 1114
QY 1621 GCTCCACCGCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGAACCCAGTG 1680
DB AlaProProAlaAlaLeuAsnProIleThrLeuLeuProGluGluThrAspGluProVal 1134
QY 1681 ACTCATGATGCCATCAACTTATGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 1740
DB ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 1154
QY 1741 ATACCGCTGACTGAGAGAGTCTAACTGTTCTACTGACGGAAGCAGCTATGTGGTGAA 1800
DB IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 1174
QY 1801 GGTAAGAGGATGGCTGGGGCGGCTGTGAGCGGACCCCGACATCTGGCCGACAGCAGC 1860
DB GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpPheAspSer 1194
QY 1861 CTGCCGGAAGGAACCTTCAGCAAAAAGCTCAGCTCATGCGCCCTCAGCAAGCTTTGCGG 1920
DB LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 1214
QY 1921 CTGCCGGAAGGAATTCATATAACGACGACAGCAGGTATGCTTTTCCGACTGCA 1980
DB LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 1234
QY 1981 CACGTACATGGGGCCATCTATAAAACAAAGGGGTTGCTTACCTCAGCAGGGAGGAAATA 2040
```

Db 1235 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 1254
QY 2041 AAGAACAAAGAGAAATTTCTAAGCCTATTAGAAGCCGTACATTATTACCAAAAAGGCTAGCT 2100
Db 1255 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 1274
QY 2101 ATTATACATCTCTCGACATCATGAAGCTAAAGATCTCATATCCAGAGGAAACCAAGATG 2160
Db 1275 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 1294
QY 2161 GCTCACCGGTTCGACAGCCAGCCAGCGGTCTTAACCTTCGCTTATATATGAATG 2220
Db 1295 AlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThr 1314
QY 2221 CCCAAAGCCCCAGAACCCAGACGACAGTACACCCCTAGAAAGCTGCGAAGAGATATAAAAG 2280
Db 1315 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 1334
QY 2281 ATAGACAGTCTCTGACATCCGGAAGGAGCTGTATACCTCAGATGGGAAGGAATC 2340
Db 1335 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 1354
QY 2341 CTGCCCCACAAGAGAGGTAGATATGTCCAAACATCCCTTATCATGTCTTCGAGCTACCAAG 2400
Db 1355 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 1374
QY 2401 ACTAAACACCTCGACGATGGTCAGAACATCCCTTATCATGTCTTCGAGCTACCAAG 2460
Db 1375 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 1394
QY 2461 GTGCTGACTCGGTGGTCAACATTTGTGTGCCCTCGCAGCTGGTTAATGCTTAATCTCTCC 2520
Db 1395 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 1414
QY 2521 AGATGCTCCAGGAAGACTAAGGGAGCCACCCAGCGCTACTCGGAAGTGAC 2580
Db 1415 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 1434
QY 2581 TTCACTCAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTTGTGTAGACACC 2640
Db 1435 PheThrGluValLysProAlaLysTyrGlyAsnLysIleLeuLeuValPheValAspThr 1454
QY 2641 TTTTCAGATGGGTAGGCTTATCTCTAATAAGAAAGAGACTTCAACCGTGGTGGCTAAA 2700
Db 1455 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 1474
QY 2701 AAAATACTGAAGAAATTTTCCAGATTTGGAATACCTAAGGTAAATAGGTCAGACAT 2760
Db 1475 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 1494
QY 2761 GGTCCAGCTTTTGTGGCCAGGTAAAGTCAGGAGCTGGCCAAAGATATTGGGGATTGATTGG 2820
Db 1495 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 1514
QY 2821 AAATGTCATTGTGCATACAGACCCCAAGCTCAGGACAGGTAGAGAGATGAATAGAAC 2880
Db 1515 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 1534
QY 2881 ATTAAGAGACCTTACTAATTTGACCGCGAGACTGGCGTTAATGATTGGATGCTCTC 2940
Db 1535 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu 1554
QY 2941 CTGCGCTTCTGCTTTTAGGGTTAGGAACACCCCTGGACAGTTTGGCTGACCCCTAT 3000
Db 1555 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1574
QY 3001 GAATTACTCTACGGGGACCCCGCCCATTTGGTGAATAATTCCTCTGTATACATAGTGTGAC 3060
Db 1575 GluLeuLeuTyrGlyGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp 1594
QY 3061 GTGCTGCTTTCCAGCCTTTGTTCTCTAGGCTCAAGGCACTTGAAGTGGGTGACACCGA 3120

Db 1595 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1614
QY 3121 GCGTGGAGGCAACTCCGGGAGGCTTACTCAGAGAGAGAGACTTGCAGATCCCATCGT 3180
Db 1615 AlaTrpLysGlnLeuArgGluAlaTyrSerGlyGluGlyAspLeuGlnValProHisArg 1634
QY 3181 TTCAAAGTGGAGATTCACTTACGTTAGACGCCCGTCAGGAAACCTCGAGACTCGG 3240
Db 1635 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1654
QY 3241 TGGAGGGCCCTTATCTCGTACTTTTGCACCAACCAACGGCTGTGAAAAGTCGAAGGATC 3300
Db 1655 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1674
QY 3301 TCCACCTGGATCCATGTCATCCCGTTAAACCGCGCCACCTCCCGATTCCGGGTGGAA 3360
Db 1675 SerThrTrpIleHisAlaSerHisValLysLeuAlaProProAspSerGlyTrpArg 1694
QY 3361 GCCGAAAGACTGAAAATCCCTTAACTTCGCTCCATCGCGTGGTTCCTTACTCTGTC 3420
Db 1695 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn 1714
QY 3421 AATAAC 3426
Db 1715 AsnAsn 1716
RESULT 7
Q8J4V8_9GAMR
ID Q8J4V8_9GAMR PRELIMINARY; PRT; 1718 AA.
AC Q8J4V8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gag-pol polyprotein.
OS Porcine endogenous retrovirus A.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=194958;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=22172947; PubMed=12185278;
RX Bartosch B., Weiss R.A., Takeuchi Y.;
RT "PCR-based cloning and immunocytological titration of infectious
porcine endogenous retrovirus subgroup A and B";
J. Gen. Virol. 83:2231-2240(2002).
DR EMBL; AY099323; AAM29192.1; -; Genomic_DNA.
DR HSSP; P03355; 116J.
DR SMR; Q8J4V8; 7-98.
DR MEROPS; A02.020; -.
DR CG; GO:0019028; C:viral capsid; IEA.
DR CG; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR CG; GO:0003677; F:DNA binding; IEA.
DR CG; GO:0004523; F:ribonuclease H activity; IEA.
DR CG; GO:0003723; F:RNA binding; IEA.
DR CG; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR CG; GO:0005198; F:structural molecule activity; IEA.
DR CG; GO:0006310; P:DNA recombination; IEA.
DR CG; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR CG; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR CG; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G retro matrix.
DR InterPro; IPR001995; Peptidase A2 cat.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR001878; Znf CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.

DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF00098; ZF_CCHC; 1.
DR SMART; SM00343; ZNF_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50158; ZF_CCHC; 1.
KW Polyprotein.
SQ SEQUENCE 1718 AA; 193115 MW; 4E86215286C59246 CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1718
Score: 5866.50 Matches: 1111
Percent Similarity: 98.4% Conservative: 15
Best Local Similarity: 97.1% Mismatches: 17
Query Match: 94.9% Indels: 1
DB: 2 Gaps: 1

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q8J4V8_9GAMR (1-1718)
QY 1 ATGGGTGCCACAGGGCAACAACAGATCCATGGACTACCCGAAGAACAGTTGACTTGGGA 60
DB 575 MetGlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgArgThrValAspLeuGly 594
QY 61 GTGGGACGGTAACCCACTCGTTCTGCTCATACCTGAGTGCACGACACCCCTCTTAGGT 120
DB 595 ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly 614
QY 121 AGAGACTTATTGACCAAGATGGGACACAAATTTCTTTTGAACAGGGAACCCAGAGTG 180
DB 615 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 634
QY 181 TCTGCAATATACAACTATCACTGTGTGTGACCTCCCAATTAGATGACGAATATCGACTA 240
DB 635 SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 654
QY 241 TACTCTCCCTAGTAAGCCTGATCAAAATATACAAATTCGTGTGGACAGTTCCCCCAA 300
DB 655 TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln 674
QY 301 GCCTGGGCAGAAACCGCAGGGATGGGTTGGCAAGAGAGTTCCCCCAACAAGTTATTCAA 360
DB 675 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 694
QY 361 CTGAAGGCCAGTGCCACACAGTGTCACTCAGACAGTACCCCTTGAGTAAGAAGCTCAA 420
DB 695 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg 714
QY 421 GAAGGAATTCGGCCGATGTCCAAAGATTAAATCAACAGGGCATCCTAGTTCTGTGCCAA 480
DB 715 GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 734
QY 481 TCTCCTCGGAATACCTCCCTGCTACCGGTTAGAAAGCCTGGGACTTAATGACTATCGACA 540
DB 735 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 754
QY 541 GTACAGGACTTGAGAGAGGTCAATAAACGGGTGCAGGATATACACCCCAACAGTCCCGAAC 600
DB 755 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 774
QY 601 CCTTATAACCTCTGTGTCTCCACCCACCAACGGAGCTGGTATACAGTATTGGACTTA 660
DB 775 ProTyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValLeuAspLeu 794
QY 661 AAGGATGCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTGGCCTTCGAA 720
DB 795 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 814
QY 721 TGGAGAGATCCAGGTACCGGGAAGAACCGGGCGAGCTCACCTGGAACCCGACTGCCCAAGGG 780
DB 815 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 834
QY 781 TTCAAGNACTCCCCGACCATCTTTTGACGAAGCCCTTACACAGAGACCTTGGCCAACTTCAGG 840

DB 835 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 854
QY 841 ATCCAAACACCTCTAGGTGACCCCTCCTCAGTACCTGATGACCTGCTTCTCGCGGGAGCC 900
DB 855 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 874
QY 901 ACCAAACAGGACTCTTAGAAGGCACGAAGCCTACTCTGTGGAAATTGTCTGACCTAGGC 960
DB 875 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuGluLeuSerAspLeuGly 894
QY 961 TACAGAGCCTCTGCTAAGAAGCCCGAGATTTGCGAGGAGAGAGTAACATACTTTGGGGTAC 1020
DB 895 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 914
QY 1021 AGTTTGGGGACGCGGACGATGGCTGACGAGGACGCGAAGAAACCTGTAGTCCAGATA 1080
DB 915 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 934
QY 1081 CCGGCCCAACACAGCCAAACAAATGAGAGATTTTGGGGACAGCTGGATTTTGCAGA 1140
DB 935 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 954
QY 1141 CTGTGGATCCCGGGGTTTGGACCTTAGCAGCCCGCTCTACCCGCTAACCAAGAAAA 1200
DB 955 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 974
QY 1201 GGGAAATCTCTGGGCTCTGAGCACAGAGCATTTGATGTATCAAAAAGGCCCTG 1260
DB 975 GlyGlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 994
QY 1261 CTGAGCCACTCTCTGGCCCTCCCTGAGCTAACTAAACCTTTACCTTTATGTGGAT 1320
DB 995 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 1014
QY 1321 GAGCTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGACCATGAGAGACCT 1380
DB 1015 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 1034
QY 1381 GTCCCTCTACCTGTCAAAGAGCTCGATCCTGTAGCAGTGGTGGCCCATATGCTGAAG 1440
DB 1035 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTyrProValCysLeuLys 1054
QY 1441 GCTATCCAGCTGGCCACTACTGTCGAAGACCTGACAAATTCGATTTTCGGACAGAAAT 1500
DB 1055 AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 1074
QY 1501 ATAACTGTATAGCCCCCATGTCATTGGAGAACATCGTTTCGGCAGCCCCCAGACCGATGG 1560
DB 1075 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 1094
QY 1561 ATGACCAAGCCCGCATGACCCCATCATCAAAAGCCTGCTTCTCACAGAGGGTTCACGTT 1620
DB 1095 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 1114
QY 1621 GCTCCACAGCCGCTCAACCCCTGCCACTCTCTGCTGGAAGACTGATGAACCACTG 1680
DB 1115 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 1134
QY 1681 ACTCATGATTCCTCATCACTATTGATTGAGGAGACTTGGGGTCCGCAAGGACCTTACAGAC 1740
DB 1135 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 1154
QY 1741 ATACCGTGTGAGGAAGTGTAACTCTGGTTCTACTGACGGAAGCAGACTAATGTGGTGAA 1800
DB 1155 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 1174
QY 1801 GGTAAAGGATGGCTGGGGCGGCTGTGTGACCGGACCCGACGATCTGGCGCAGAGC 1860
DB 1175 GlyLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAlaSerSer 1194
QY 1861 CTGCCGGAAGGAACCTTCAGCAAAAAAGGCTCAGCTCATGGCCCTCACCAAGCTTTGCGG 1920

Db 1195 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 1214
 Qy 1921 CTGCCCGAAGGAAATCATATAACATTTATACGACAGCAGGTATGCCCTTGGCACTGCA 1980
 Db 1215 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 1234
 Qy 1981 CACGTACATGGGCGCATCTATAACAAGAGGGGTGCTTACTCAGCGAGGAGGAATA 2040
 Db 1235 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 1254
 Qy 2041 AAGAACAAAGAGGAAATCTAAGCCTATTAGAAGCGGTACATTTACCAAAAAGGCTAGCT 2100
 Db 1255 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 1274
 Qy 2101 ATTATACATGCTCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATG 2160
 Db 1275 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 1294
 Qy 2161 GCTCACCAGGTTGCCAAGCAGCGCCAGCGGTGTTAACTTCTGCCTTAATAAGAAATG 2220
 Db 1295 AlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThr 1314
 Qy 2221 CCCAAAGCCCGAAGAACCCAGACGACGTACACCTTAGAAGACTGCGAAGAGATAAAAAG 2280
 Db 1315 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 1334
 Qy 2281 ATAGACCATGTTCTCGAGACTCCGGAAGGACCTGCTATATCCTCAGATGGGAAGAAATC 2340
 Db 1335 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 1354
 Qy 2341 CTGCCCCAAGAAAGGGTTAGAATATGTCCAAACAGATACATCTCTAACCCACCTAGGA 2400
 Db 1355 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 1374
 Qy 2401 ACTAAACACCTCGACGAGTGGTCAGAACATCCCTTATCATGTTCTGAGCTACACGGA 2460
 Db 1375 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 1394
 Qy 2461 GTGGCTGACTCGGTGGTCAAAACATTTGTGTGCTCCAGCTGGTAAATGCTAATCTCTCC 2520
 Db 1395 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 1414
 Qy 2521 AGAATGCTCCAGGAAAGAGACTAAGGGGAAGCCACCAGCGCTCACTGGGAAGTGGAC 2580
 Db 1415 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 1434
 Qy 2581 TTCACGTAGCTAAGCCGGCTAAATACGGAACAAATACCTATTGTTTGTGTAGACACC 2640
 Db 1435 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 1454
 Qy 2641 TTTTTCAGGATGGGTAGAGGCTTATCCTTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 2700
 Db 1455 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 1474
 Qy 2701 AAAATACCTGAAGAAATTTTCCAAAGTTTGGAAATACCTAAGGTAAATAGGCTCAGACAAT 2760
 Db 1475 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 1494
 Qy 2761 GSTCCAGCTTTGTGCGCAGGTAAAGTACGAGGACTGCGCAAGATATTGGGGATTGATGG 2820
 Db 1495 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 1514
 Qy 2821 AAATCTGATTTGTGCATACAGACCCAAAGCTCAGGACAGGTAGAGAGATGAATAGAACCC 2880
 Db 1515 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 1534
 Qy 2881 ATTAAGAGACCTTACTAAATTCAGCGCGAGACTGGCGCTTAATGATGATGCTCTC 2940
 Db 1535 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeu 1554
 Qy 2941 CTGCCCTTTGTGCTTTTAGGGTTAGGAACACCCCTGGACAGTTTGGCTGACCCCTAT 3000
 Db 1555 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1574

Qy 3001 GAATTACTTACGGGGGACCCCCCATTTGGTAGAAATTTGCTTCTGTACATAGTCTGAC 3060
 Db 1575 GluLeuLeuTyrGlyGlyProProLeuAlaGluIleAlaPheAlaHisSerAlaAsp 1594
 Qy 3061 GTGCTGCTTTCCAGCCTTTGTTCTTAGGCTCAAGCCTTGAAGTGGGTGAGACACGA 3120
 Db 1595 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1614
 Qy 3121 GCGTGGAGGCAACTCCGGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCCATCGT 3180
 Db 1615 AlaTrpLysGlnLeuArgGluAlaTyrSer---GlyGlyAspLeuGlnValProHisArg 1633
 Qy 3181 TTCCAAGTGGGAGATTCACTTACGTTAGACGCCACCGTGCAGGAAACCTCGAGACTCGG 3240
 Db 1634 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1653
 Qy 3241 TGGAGGGCCCTTATCTCGTACTTTTGACACACCAACCGGTGTGAAGTGAAGGAATC 3300
 Db 1654 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1673
 Qy 3301 TCCACCTGGATCATCATCCAGTTCAGTTAAACCGCGCCACCTCCCGATTGGGGTGGAAA 3360
 Db 1674 ProThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1693
 Qy 3361 GCCGAAAGACTGAAATCCCTTAAGCTTCGCCTCCATCGCGTGTCTTACTCTGTCT 3420
 Db 1694 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1713
 Qy 3421 AATAACCTCTCA 3432
 Db 1714 AsnAsnSerSer 1717
 RESULT 8
 Q9Q1X5_9GAMR
 ID Q9Q1X5_9GAMR PRELIMINARY; PRT; 2376 AA.
 AC Q9Q1X5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Type C proviral gag, pol and env genes and LTR (class B, clone 33).
 DE Porcine endogenous retrovirus.
 OS Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Gammaretrovirus; 1-Mammalian type C virus group.
 OX NCBI_TaxID=61673;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=type C;
 RX MEDLINE=20219394; PubMed=10756014;
 RX DOI=10.1128/JVI.74.9.4028-4038.2000;
 RA Czauderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;
 RT "Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells";
 RL J. Virol. 74:4028-4038(2000).
 DR EMBL; AJ133816; CAB65339.1; -; Genomic_DNA.
 DR HSSP; P03385; 1MOP.
 DR SNR; Q9Q1X5; 7-98, 472-518, 2225-2277.
 DR MEROPS; A02.020; -.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR GO; GO:0019068; P:viral assembly; IEA.
 DR InterPro; IPR002050; Env_polyprotein.
 DR InterPro; IPR008981; FmuLVrecept-bind.
 DR InterPro; IPR003036; Gag_p30.

```
DR InterPro; IPR000840; G_retro_matrix.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR001969; Pept_Asp_AS_.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF00429; TLV_coat; 1.
DR Pfam; PF00098; zf_CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS01175; ASP_PROT_RETROV; 1.
DR PROSITE; PS50879; RNASE_H_1.
DR PROSITE; PS50158; ZF_CCHC; 1.
FT CHAIN 1 524 Gag protein.
FT CHAIN 525 1761 Pol protein.
FT CHAIN 1720 2376 Env protein.
SQ SEQUENCE 2376 AA; 266625 MW; 37DD3D3099DA4AF CRC64;

Alignment Scores:
Pred. No.: 0 Length: 2376
Score: 5864.00 Matches: 1109
Percent Similarity: 98.5% Conservative: 16
Best Local Similarity: 97.1% Mismatches: 17
Query Match: 94.8% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q90IX5_9GAMR (1-2376)
Qy 1 ATGGGTGCCAGCGCAACACAGTATCATGTGACTACCCGAGAACAGTGTGACTTGGGA 60
Db 574 MetGlyAlaThrGlyGlnArgGlnTyProTrpThrArgArgThrValAspLeuAla 593
Qy 61 GTGGAGCGGTAAACCACTCGTTTCTGGTCATACCTGAGTCCGACGACCCCTCTTAGGT 120
Db 594 ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly 613
Qy 121 AGAGACTTATTGACCAAGATGGGACCAATTTCTTTGAAACAGGGAACCAAGAGTG 180
Db 614 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 633
Qy 181 TCTGCAATAACAAACCTATCACTGTGTGTGACCTCCCAATTAGATGACGAATATCGACTA 240
Db 634 SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyArgLeu 653
Qy 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAATTTCTGGTTGGAAACAGTTTCCCAA 300
Db 654 TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln 673
Qy 301 CCTGGGCGAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCGCCCAACAGTTATTCAA 360
Db 674 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 693
Qy 361 CTGAAGGCCAGTGCCACACCACTGTCAGTCAGACAGTACCCCTTGAGTAAAGAAGCTCAA 420
Db 694 LeuLysAlaSerAlaThrProValSerValArgGlnTyProLeuSerArgGluAlaArg 713
Qy 421 GAAGGAATTCGGCCGATGTCCAAAGATTAAATCCAACAGGGCATCTTAGTTCTGTCCAA 480
Db 714 GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 733
Qy 481 TCTCCCTGGAACTACCTCCCTGCTACCGTTAGAACGCTGGACTTAATGACTATCGACCA 540
Db 734 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyArgPro 753
Qy 541 GTACAGGACTTGAGAGAGGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAAC 600
```

```
Db 754 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 773
Qy 601 CCTTATAACCTCTTGTGTGCTCTCCACCCCAAGGAGCTGGTATACAGTATTGCAGCTTA 660
Db 774 ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyThrValLeuAspLeu 793
Qy 661 AAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAA 720
Db 794 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 813
Qy 721 TGGAGAGATCCAGGTACGGGAAGAACCGGCAGCTCACCTGGACCCCACTGCCCAAGGG 780
Db 814 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 833
Qy 781 TTCAGAAGCTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTCAGG 840
Db 834 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 853
Qy 841 ATCCAACACCTCAGGTGACCCCTCCTCAGTACGTGATGACCTGCTTCTGGCGGAGCC 900
Db 854 IleGlnHisProGlnValThrLeuLeuGlnTyThrValAspAspLeuLeuAlaGlyAla 873
Qy 901 ACCAAACAGGACTGCTTTAGAGGACCAAGACACTACTGTGGAATTGTCTGACTAGGC 960
Db 874 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 893
Qy 961 TAGAGACCTCTGCTAAGAGGCCAGATTTCAGGAGAGAGGTAACATACCTTGGGGTAC 1020
Db 894 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyLeuGlyTy 913
Qy 1021 AGTTTGGCGGACGGCGACGATGGCTGACGGAGGACGGAAGAAACCTGTAGTCCAGATA 1080
Db 914 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 933
Qy 1081 CCGGCCCCAACCAAGCAAAATGAGAGAGTTCCTTGGGGACAGCTGGATTTCGAGA 1140
Db 934 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 953
Qy 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTACCCGCTAACCCAAAGAAAAA 1200
Db 954 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyProLeuThrLysGluLys 973
Qy 1201 GGGGAATTCCTCGGGCTCCTGAGCACAGAAAGCATTTGATGCTATCAAAAAGGCCCTG 1260
Db 974 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 993
Qy 1261 CTGAGCGACACTGCTCTGGCCCTCCCTGACGTAACCTAAACCCCTTACCCCTTATGTGAT 1320
Db 994 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyValAsp 1013
Qy 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGACCATGGAGAGACCT 1380
Db 1014 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgPro 1033
Qy 1381 GTCCCTCACTCTGTCAAAGAGCTCGATCCTGTAGCAGTGGTGGCCCATATGCTCGAAG 1440
Db 1034 ValAlaTyLeuSerLysLysLeuAspProValAlaSerGlyTyProValCysLeuLys 1053
Qy 1441 GCTATCGCAGCTGTGGCCATCTGGTCAAGGACGCTGACAAATTCATCTTGGGACAGAA 1500
Db 1054 AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 1073
Qy 1501 ATAACGTAAATAGCCCCCATGCAATTGGAGAACATCGTTCCGGCAGCCCCCAGACCGTGG 1560
Db 1074 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 1093
Qy 1561 ATGACCAACGCCCGCATGACCCACTATCAAGCCCTGCTTCTCACAGAGAGGCTCACGTTTC 1620
Db 1094 MetThrAsnAlaArgMetThrHisTyGlnSerLeuLeuLeuThrGluArgValThrPhe 1113
Qy 1621 GCTCCACAGCGCTCTCAACCCCTGCCACTCTTCTGCTGGAAGAGACGTGTAACAGTG 1680
Db 1114 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 1133
```


QY 1691 ACTCATGATGCCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 1740
Db ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 1153
QY 1741 ATACCGCTGACTGGAGAGTCTAACTGGTTCCTACTGACGGAAGAGCTATGTGTGAA 1800
Db IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 1173
QY 1801 GGTAAAGAGATGGCTGGGGCGCGGTGTGGACGGGACCCGACGATCTGGGCCAGCAGC 1860
Db GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrHisThrIleTrpAlaSerSer 1193
QY 1861 CTGCGGGAAGAACTTTCAGCAAAAGGCTGAGCTCATGGCCCTCACCAAGACTTTGGCG 1920
Db LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 1213
QY 1921 CTGCGCGAAGGAATCCATAAACAATTTATAGGACAGCAGGTATGCCCTTTGGCACTGCA 1980
Db LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 1233
QY 1981 CACGTACATGGGGCCATCTATAAACAAGGGGTTGCTTACCTCAGCAGGAGGAGAAATA 2040
Db HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 1253
QY 2041 AAGAACAAAGAGAAATTTCTAAGCCTATTAGAAGCCGTACATTTACCACAAAAGGCTAGCT 2100
Db LysAsnLysGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 1273
QY 2101 ATTATACACTCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATG 2160
Db IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 1293
QY 2161 GCTCACCGGGTTGCCAAGCAGCAGCCAGGGTGTAACTCTGCGCTTAATAAGAAATG 2220
Db AlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThr 1313
QY 2221 CCCAAGCCCAAGAACCCAGACGACGACAGTACACCCCTAGAAAGACTGGCAAGAGATAAAAAAG 2280
Db ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 1333
QY 2281 ATAGACAGATCTCTGAGACTCCGGAAGGACCTGTATACCTCAGATGGGAAGAAATC 2340
Db IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 1353
QY 2341 CTGCCCCACAAGAGAGGTTAGAAATATGTCACAGATACATCTCTAAACCCACTAGGA 2400
Db LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 1373
QY 2401 ACTAAACACCTGCAGCAGTGGTTCAGAACATCCCTTATCATGTTCTGAGGCTACCAAGGA 2460
Db ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 1393
QY 2461 GTGCTGACTCGGTGGTCAAACTATGTGTGCCCTGCCAGCTGGTTAATCTTAATCTCTCC 2520
Db ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 1413
QY 2521 AGAATGCTCCAGGGAAGACTAAGGGAGCCACCCAGCGCTCACTGGAGTGCAC 2580
Db ArgIleProProGlyLysArgLeuArgLysSerHisProGlyAlaHisTrpGluValAsp 1433
QY 2581 TTCACTGAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGGTTTGTGTAGACACC 2640
Db PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 1453
QY 2641 TTTTCAGATGGGTAGAGCTTATCTCTACTAAGAAAGAGACTTCAACCGGTGGTGGTAA 2700
Db PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 1473
QY 2701 AAAATACCTGAAGAAATTTTCCAGATTTGGATACCTAAGGTAAATAGGGTCAGACAAT 2760
Db LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 1493

QY 2761 GGTCAGCTTTTGTGGCCAGGTAAAGTCAGGACTGGCCAAAGATATTGGGATTCATTTGG 2820
Db GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 1513
QY 2821 AAACCTGATTTGTGATACAGACCCCAAGACTCAGGAAGGTAGAGAGATGAATAGAAC 2880
Db LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 1533
QY 2881 ATTTAAAGAGACCTTTACTTAAATTCACCGCGAGAGCTGCGCTTAATGATTGGATGCTCTC 2940
Db IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu 1553
QY 2941 CTGCCCCTTTGTGCTTTTGTAGGTTAGGAACACCCCTGGACAGTTTGGGCTCAGCCCTAT 3000
Db LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1573
QY 3001 GAATTACTCTACGGGGACCCCTCCCTATTTGGTAGAAAATTTGTTGATCATAGTCTGAC 3060
Db GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1593
QY 3061 GTGCTGCTTTCCAGACCTTTGTTCTTAGGCTCAAGCACCTTGAGTGGGTGAGACAACGA 3120
Db ValLeuLeuSerGlnProLeuPheSerArgLysAlaLeuGluTrpValArgGlnArg 1613
QY 3121 GCGTGGAGGCACTCCGGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCCATCGT 3180
Db AlaTrpLysGlnLeuArgGluAlaTyrSerGlyGluGlyAspLeuGlnValProHisArg 1633
QY 3181 TTCCAAGTGGAGATTTCAGTCTACGTTAGACGCCACCGTCGAGAAACCTCGAGACTCGG 3240
Db PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1653
QY 3241 TGAAGGCGCTTATCTCTGCTTTCACACACCAACCGCTGTGAAGTCCGAAGGAATC 3300
Db TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1673
QY 3301 TCACCTGGATTCATGTCATCCACGTTTAAACCCGCGCCACCTCCCGATTCGGGTGGA 3360
Db SerThrTrpIleHisThrSerHisValLysLeuAlaProProAspSerGlyTrpArg 1693
QY 3361 GCGGAAAGACTGAAATATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTC 3420
Db AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn 1713
QY 3421 AATAAC 3426
Db 1714 AsnAsn 1715
RESULT 9
Q8UMP5_9GAMR
ID Q8UMP5_9GAMR PRELIMINARY; PRT; 1146 AA.
AC Q8UMP5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol.
OS Porcine endogenous retrovirus.
OC Viruses; Retrov. viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21568379; PubMed=11711616;
DOI=10.1128/JVI.75.24.12252-12265.2001;
RA Herring C., Quinn G., Bower R., Parsons N., Logan N.A., Brawley A.,
RA Elsom K., Whittam A., Fernandez-Suarez X.M., Cunningham D.,
RA Onions D., Langford G., Scobie I.,
RT "Mapping full-length porcine endogenous retroviruses in a large white
pig".
RL J. Virol. 75:12252-12265(2001).
DR EMBL; AY056035; AAL38193.1; -; Genomic_DNA.
DR HSP; P03355; 116J.
DR MEROPS; A02.020; -.

DR GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO:0003677; F:DNA binding; IEA.
DR GO:0004523; F:ribonuclease H activity; IEA.
DR GO:0003723; F:RNA binding; IEA.
DR GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO:0006310; P:DNA recombination; IEA.
DR GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR001995; Peptidase_A2_cat.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00075; RNaseH; 1.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00077; RVP; 1.
DR Pfam: PF00078; RVT_1; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS0879; RNASE H; 1.
SQ SEQUENCE 1146 AA; 128417 MW; 1244B29FC5C270DE CRC64;

Alignment Scores:

Pred. No.:	0	Length:	1146
Score:	5863.00	Matches:	1109
Percent Similarity:	98.6%	Conservative:	17
Best Local Similarity:	97.1%	Mismatches:	16
Query Match:	94.8%	Indels:	0
DB:	2	Gaps:	0

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q8UMP5_9GAMR (1-1146)

QY	1	ATGGTGCCACAGGCAACACAGTATCCATGGACTACCCGGAAGAACAGTTGACTTGGGA	60
DB	1	MetGlyAlaThrGlyGlnArgGlnTyrProThrThrArgArgThrValAspLeuGly	20
QY	61	GTGGACGGGTAAACCACCTCGTTTCTGGTCATACCTGAGTGGCCAGCACCCCTCTTAGGT	120
DB	21	ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly	40
QY	121	AGAGACTTATTGACCAAGATGGGACCAAAATTTCTTTGAAACAGGGAACCAAGAGTG	180
DB	41	ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal	60
QY	181	TCTGCAATAAACAACCTATCACTGTGTGACCTCCAAATTAGATGACGAATATCGACTA	240
DB	61	SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu	80
QY	241	TACTCTCCCTAGTAAAGCCTGATCAAAATATACAATTCGTGTTGGACAGTTTCCCCAA	300
DB	81	TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln	100
QY	301	GCCTGGGAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCCAAGTTATTCAA	360
DB	101	AlaTrpAlaGluThrAlaGlyIleGlyLeuAlaLysGlnValProProGlnValIleGln	120
QY	361	CTGAAGCCAGTGCACACCAAGTGTCACTGACAGACTACCCCTTGAGTAAAGAAGCTCAA	420
DB	121	LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg	140
QY	421	GAAGGAATTCGGCCGTCATGTCCAAGATTAAATCCAACAGGGCATCTAGTTCTCTGTCAA	480
DB	141	GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuAlaProValGln	160
QY	481	TCTCCTGGGAATATCTCCCTGCTACCGTTAGAAAGCCTGGGACTAATGACTATCGACCA	540
DB	161	SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro	180
QY	541	GTACAGGACTTGGAGAGGTCAATAACGGGTGCAGGATATACCCCAACAGTCCCGAAC	600
DB	181	ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn	200
QY	601	CTTTATAACCTCTGTGTGCTCTCCACCCCAACCGAGCTGGTATACAGTATTGGACTTA	660
DB	201	ProTyrAsnLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeu	220

QY	661	AAGGATCGCTTCTTCTGCCTGAGATTACCCCACTAGCCAACCACTTTTTCCTTCGAA	720
DB	221	LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu	240
QY	721	TGGAGAGATCCAGGTACGGGAAGAACCGGCAGCTCACTCGACCCGACTGCCCAAGGG	780
DB	241	TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly	260
QY	781	TTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGG	840
DB	261	PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg	280
QY	841	ATCCAACACCTCAGGTGACCTCTCCTCAGTACGTGATGACCTGCTTCTCGCGGAGCC	900
DB	281	IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla	300
QY	901	ACCAACAGGACTGCTTAGAAGGCACGAAGCACTACTGTGGAATTGTGTGACTAGGC	960
DB	301	ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuGluLeuSerAspLeuGly	320
QY	961	TACAGACCTCTGCTAAGAAGCCACAGATTTGCAGGAGAGAGGTAAACATCTTGGGTAC	1020
DB	321	TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr	340
QY	1021	AGTTTGGCGACGGCGAGCGATGGCTGACGAGGCACGGAAGAAACTGTAGTCCAGATA	1080
DB	341	SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysThrValValGlnIle	360
QY	1081	CCGGCCCCAACACAGGCCAAACAAATGAGAGAGTTTTTTGGGGACAGCTGGATTTGCAGA	1140
DB	361	ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg	380
QY	1141	CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAAA	1200
DB	381	LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys	400
QY	1201	GGGGAATTCTCTGGGCTCTGAGCACCAAGAGCATTTGATGTATCAAAAGGCGCCTG	1260
DB	401	GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu	420
QY	1261	CTGAGCACACCTGCTCGGCCCTCCCTGACGTAACTAAACCCCTTTACCTTTATGTGAT	1320
DB	421	LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp	440
QY	1321	GAGGCTAAGGAGTAGCCCGGGAGTTTAAACCAACCCCTAGCACCATTGAGAGACCT	1380
DB	441	GluArgLysGlyValAlaAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgPro	460
QY	1381	GTCCCTTACCTGTCAAAGAAGCTCGATCTGTAGCCAGTGGTTGGCCCATATGCTGAAG	1440
DB	461	ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys	480
QY	1441	GCTATCAGCTGTGGCCATACCTGGTCAAGGACGCTGACAAATTCACCTTTGGGACAGAA	1500
DB	481	AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuArgGlnAsn	500
QY	1501	ATAACTGTATATAGCCCCCATGATTGGAGAACATCGTTTCGGCAGCCCCAGACCGATGG	1560
DB	501	IleThrValIleAlaProHisAlaLeuGluAlaAsnIleValArgGlnProProAspArgTrp	520
QY	1561	ATGACCAACGCCCATGACCCACTATCAAGCCCTGCTTCTCACAGAGAGGTCACCGTTC	1620
DB	521	MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe	540
QY	1621	GCTCCACGACCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGAACCACTG	1680
DB	541	AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal	560
QY	1681	ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGTCCCGAAGGACCTTACAGAC	1740
DB	561	ThrHisAspCysHisGlnLeuLeuIleGluThrGlyValArgLysAspLeuThrAsp	580

QY 1741 ATACCGCTGACTGGAGAAGTCTTAACCTGGTTCACCTGACGGAAGACGCTATGTGGTGA 1800
DB 581 ILeProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
QY 1801 GGTAAAGAGATGGCTGGGGCGCGTGTGTGACCGGACCCGACGATCTGGGCCAGCAGC 1860
DB 601 GlyLysArgMetAlaGlyAlaValValaspGlyThrArgThrIleTrpAlaSerSer 620
QY 1861 CTGCCGGAAGAACTTCAGCAAAAGGCTGAGCTCATGGCCCTCAGCAAGCTTTGCGG 1920
DB 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 1921 CTGCCGGAAGGAAATCCATAAATTTATACGACAGCAGGATGCTTTGCCAGCTGCA 1980
DB 641 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
QY 1981 CACGTACATGGGGCCATCTATAAACAAGGGGTTGCTTACCTCAGCAGGAGGAGAAATA 2040
DB 661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 2041 AAGAACAAAGAGGAAATTTCTAAGCCTATTAGAAGCCGTACATTTACCAAAAGGCTAGCT 2100
DB 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 700
QY 2101 ATTATACACTGCTCTGGACATCAGAAAGCTTAAGATCTCATATCCAGAGGAAACAGATG 2160
DB 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
QY 2161 GCTCACCAGGTTGTCAGAGCAGCCAGCGTGTAACTTCTGCTTATTAATAGAAATG 2220
DB 721 AlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuLeuProIleIleGluThr 740
QY 2221 CCCAAAGCCCGAGAACCCAGACGACGATACACCTTAGAAGACTGGCAAGAGATAAAAAG 2280
DB 741 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
QY 2281 ATAGACACTGCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATC 2340
DB 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 780
QY 2341 CTGCCCCCAAGAAAGGTTAGATATGTCACAGATACATCGCTAAACCCACCTAGGA 2400
DB 781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 2401 ACTAAACACCTGCAGCAGTGTGGTCAGAACATCCCTTATCATGTTCTGAGGCTACCAAG 2460
DB 801 ThrLysHisLeuGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
QY 2461 GTGGCTGACTCGGTGGTCAACATTTGTGTGCTCCGACGCTGGTTAATGCTAATCTCTCC 2520
DB 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY 2521 AGAATGCTCCAGGGAAGACTAAAGGGGAAGCCACCGCGCTCACTGGGAGTGCAC 2580
DB 841 ArgIleProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
QY 2581 TTCACTGAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTTGTAGACACC 2640
DB 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
QY 2641 TTTTCAGATGGGTAGAGCTTATCTCTACTAAGAAAGACTTCAACCGTGGTGGCTAAA 2700
DB 881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
QY 2701 AAAATCTGGAGAAATTTTCCAGATTTGGATACCTAAGCTTAAGGTCAGCAAT 2760
DB 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
QY 2761 GGTCACAGCTTTGTTGCCAGGTAAAGTCAGGACTGGCCCAAGATATTGGGGATTGATTGG 2820
DB 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY 2821 AAATGCAATGTGCATACAGACCCCAAGCTCAGGACAGGTAGAGAGGATGAATAGAAC 2880

DB 941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY 2881 ATTAAGAGACCCCTTACTTAAATTCACCGGAGACTGGCGTTAATGATTGGATAGCTCTC 2940
DB 961 IleLysGluThrLeuThrLysLeuThrGluThrGlyIleAsnAspTrpIleAlaLeu 980
QY 2941 CTGCCCTTTGTGCTTTTGTAGGTTAGGAACACCCCTCGACAGTTTGGGCTGACCCCTAT 3000
DB 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
QY 3001 GAATTACTCTACGGGGACCCCCCATTTGGTAGAAATTTGCTTCTGTACATAGTCTGAC 3060
DB 1001 GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
QY 3061 GTGCTGCTTTCCAGCCCTTTGTTCTTAGGCTCAAGGCACCTTGAGTGGGTGAGACAACGA 3120
DB 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY 3121 GCGTGGAGGCAACTCCGGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCCATCGT 3180
DB 1041 AlaTrpLysGlnLeuArgGluAlaTyrSerGlyGluGlyAspLeuGlnValProHisArg 1060
QY 3181 TTCCAAGTGGAGATTTCAGTCTAGCTTAGACGCCACCGTCGAGGAAACCTTCGAGACTCGG 3240
DB 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
QY 3241 TGAAGGCGCTTATCTGCTACTTTTCCACACACCGCTGGAAGTGAAGTGAAGGATC 3300
DB 1081 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
QY 3301 TCCACCTGGATCCATGTCATCCACGTTTAAACCCGCGCCACCTCCCGATTTCGGGTGGA 3360
DB 1101 SerThrTrpIleHisAlaSerHisValLysLeuAlaProProProAspSerGlyTrpArg 1120
QY 3361 GCGGAAAGACTGAAATTCCTTAAGCTTCGCTCCATCGCTGGTGGTTCCTTACTCTGTC 3420
DB 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn 1140
QY 3421 AATAAC 3426
DB 1141 AsnAsn 1142

RESULT 10
Q9RL6_9GAMR
ID Q9RL6_9GAMR PRELIMINARY; PRT; 1195 AA.
AC Q9RL6_9GAMR
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Namespol;
OS Porcine endogenous type C retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammaetrovirus.
OX NCBI_TaxID=70540;
RN [1]
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=21256017; PubMed=11356953;
RX DOI=10.1128/JVI.75.12.5465-5472.2001;
RA Krach U., Fischer N., Czaderna F., Toenjes R.R.;
RT "Comparison of replication-competent molecular clones of porcine
endogenous retrovirus class A and class B derived from pig and human
cells";
RL J. Virol. 75:5465-5472(2001).
CC -/- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AJ293657; CAC39626.1; -; Genomic DNA.
DR HSSP; P03355; 116J.
DR MSROPS; A02_020; -;
DR GO; GO:004390; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.

DR GO: 0003723; F: RNA binding; IEA.
DR GO: 0003964; F: RNA-directed DNA polymerase activity; IEA.
DR GO: 0006310; P: DNA recombination; IEA.
DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
DR GO: 0006278; P: RNA-dependent DNA replication; IEA.
DR InterPro: IPR001969; Pept Asp AS.
DR InterPro: IPR009007; Pept Aspartic cat.
DR InterPro: IPR001995; Peptidase_A2_cat.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR012337; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00075; RNaseH; 1.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00077; RVP; 1.
DR Pfam: PF00078; RVT; 1; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS00994; INTEGRASE; 1.
DR PROSITE: PS06879; RNASE_H; 1.
DR PROSITE: PS08878; RT_POL; 1.
KW Aspartyl protease; Hydrolase; Protease.
FT NON_TER 1
SQ SEQUENCE 1195 AA; 133770 MW; 999F1524DE0BFFA4 CRC64;

Alignment Scores:

Pred. No.:	0	Length:	1195
Score:	5861.00	Matches:	1108
Percent Similarity:	98.5%	Conservative:	17
Best Local Similarity:	97.0%	Mismatches:	17
Query Match:	94.8%	Indels:	0
DB:	2	Gaps:	0

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q90RL6_9GAMR (1-1195)

QY	1	ATGGGTGCACAGGGCAACACAGTATCCATGGACTACCCGAGAACAGTTGACTTTGGGA	60
DB	50	MetGlyAlaThrGlyGlnArgGlnTyrProIleThrThrArgThrValAspLeuAla	69
QY	61	GTGGACGGGTAACCCACTCGTTTCTGGTCATACCTGAGTGCACAGCACCCCTCTTAGGT	120
DB	70	ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly	89
QY	121	AGAGACTTATGCACCAAGATGGGACCAAAATTTCTTTTGAACAAGGAAACCAAGATG	180
DB	90	ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal	109
QY	181	TCTGCAAAATAACAACTTATCACTGTGTGACCTCCCAATTAGATGACGAATATCGACTA	240
DB	110	SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspGluTyrArgLeu	129
QY	241	TACTCTCCCTAGTAAGCCCTGATCAAAATATACAAATTTCTGGTTGGAAACAGTTTCCCAA	300
DB	130	TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln	149
QY	301	GCCTGGGAGAAACCGCAGGAGATGGTTGGCAAGAGAGTTCCCCCAAGATTATTCAA	360
DB	150	AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln	169
QY	361	CTGAAGGCCAGTGCCACACCACTGTCAGTCAGACAGTACCCCTTCAGTAAAGAGCTCAA	420
DB	170	LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg	189
QY	421	GAAGAAATTCGGCCGCGCATGTCCAAAGATTAATCCAAACAGGGGCATCTAGTTCTCTGTC	480
DB	190	GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln	209
QY	481	TCTCCCTGGATACTCCCTGCTACCGTTAGAAAGCCTGGGACTAATGACTATCGACCA	540
DB	210	SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro	229
QY	541	GTACAGGACTTGAGAGAGTCAATAAACGGGTGCAGGATATACACCAACAGTCCCGAAC	600

DB	230	ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn	249
QY	601	CCTTATACCTCTTGTGTGTCTCCACCCCAACGAGAGCTGGTATACAGTATTGGACTTA	660
DB	250	ProTyrAsnLeuLeuSerAlaLeuProGluArgAsnTrpTyrThrValLeuAspLeu	269
QY	661	AAGGATGCTCTTCTCTCGCTGAGATTACACCCACTAGCCCAACCACTTTTTCCTTCGAA	720
DB	270	LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu	289
QY	721	TGGAGAGATCCAGGTACGGGAAGAACCGGCGAGCTCACCTGGACCCGACTCCCCCAAGGG	780
DB	290	TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly	309
QY	781	TTCAGAATCTCCCGACCATCTTTGACAGCCCTCACAGAGACCTGGCCCACTTCAGG	840
DB	310	PheLysAsnSerProThrPhePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg	329
QY	841	ATCCAACACCTCAGGTGACCTCCTCCAGTACCTGGATGACCTGCTCTCGCGGGAGCC	900
DB	330	IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla	349
QY	901	ACCAACACGAGTCTCTTAGAAGCAGCAAGGCACTACTGCTGGAATTGCTCGACTAGGC	960
DB	350	ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly	369
QY	961	TACAGAGCTCTGTAAGAGGCCAGATTTCGAGGAGAGAGGTAACATATCTTGGGGTAC	1020
DB	370	TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr	389
QY	1021	AGTTTGGGGACGGCGAGCGATGCTGACGAGGACGACGAGAAACCTGTAGTCCAGATA	1080
DB	390	SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle	409
QY	1081	CCGGCCCCCAACACAGCCCAACAAATGAGAGAGATTTTGGGGACAGCTGGATTTCGAGA	1140
DB	410	ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg	429
QY	1141	CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAA	1200
DB	430	LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLys	449
QY	1201	GGGGAATTTCTCTGGCTCCTGACACACAGAGGATTTGATGCTATCAAAAAGGCCCTG	1260
DB	450	GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu	469
QY	1261	CTGAGCGCACCTGCTCTGGCCCTCCCTGACATAACTAAACCTTTTACCTTTATGTGGAT	1320
DB	470	LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp	489
QY	1321	GAGGTAAAGGAGTAGCCCGGGAGTTTAAACCCAAACCCCTAGGACCATGAGAGACCT	1380
DB	490	GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro	509
QY	1381	GTCCCTCACTCTCAAGAGAGCTCGATCCTGTAGCAGTGTGGCCCATATGCTCGAAG	1440
DB	510	ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys	529
QY	1441	GCTATCGCAGCTGTGGCCCATCTGTCAGGACCGCTGCAAAATTTGACTTTTGGGACAGAT	1500
DB	530	AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn	549
QY	1501	ATAACTGTAAATAGCCCCCATGCAATTGGAGAACATCGTTTCGGGAGCCCCCAGACCGATGG	1560
DB	550	IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp	569
QY	1561	ATGACCAACGCCCGCTGACCCACTATCAAGAGCTGCTTCTCACAGAGAGGTTCACGTTTC	1620
DB	570	MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe	589
QY	1621	GCTCCACGAGCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGTGAACCACTG	1680
DB	590	AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal	609

QY 1681 ACTCATGATGCCATCAACTATTGATTGAGGAGACTGGGTCCGCAAGACCTTTACAGAC 1740
DB ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 629
QY 1741 ATACCGCTGACTGGAGAGTCTTAACCTGGTTCACTGACGGAAGCAGTATGTTGGTGA 1800
DB IleProLeuThrGlyValLeuLeuThrTrpPheThrAspGlySerSerTyrValValGlu 649
QY 1801 GGTAAAGAGCATGGTGGCGCGGTGGTGCACGGGACCCGACAGCATCTGGCGCCAGCAGC 1860
DB GlyLysLysMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 669
QY 1861 CTGCGCGAAGAACTTCAGCAAAAAGGCTGAGTCACTGGCCCTCAGCAAGCTTTGCGG 1920
DB LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 689
QY 1921 CTGCGCGAAGGAAATCCATAAATTTATACGACAGCAGGATGCTTTGCCACTGCA 1980
DB LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 709
QY 1981 CACGTACATGGCGCATCTATAAACAAGGGGTTGCTTACCTCAGCAGGAGGAGAAATA 2040
DB HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 729
QY 2041 AAGAAACAAAGAGAAATTTAAGCCTATTAGAAGCCGTACATTTACCAAAAAGGCTAGCT 2100
DB LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 749
QY 2101 ATTATACATCTCTCGACATCAGAAAGCTTAAGATCTCATATCCAGAGGAAACCATG 2160
DB IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 769
QY 2161 GCTCACCGGTTGCCAAGCAGGACCCAGGCTGTTAAACCTTCGCTATAATAAGAAATG 2220
DB AlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuLeuProIleIleGluThr 789
QY 2221 CCCAAAGCCCGAAGCAGACGACGTACACCTTAGAAGACTGGCAAGAGATAAAAAAG 2280
DB ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 809
QY 2281 ATAGACCATGCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATC 2340
DB IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 829
QY 2341 CTGCCCCACAAGAAAGGTTAGATATGTCACAGATACATGCTTAACCCACCTAGGA 2400
DB LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 849
QY 2401 ACTAAACACCTGCAGCAGTGGTCAGAACATCCCTTTATCATGTTCTGAGGCTACCAAGGA 2460
DB ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuValArgLeuProGly 869
QY 2461 GTGCTGACTCGTGGTCAACATTTGTGTCCTCCGACCTGGTAAATGTAATCTCTTC 2520
DB ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 889
QY 2521 AGAATGCTCCAGGGAAGACTAAGGGGAAGCCACCGCGCTCACTGGGAGTGCAC 2580
DB ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 909
QY 2581 TTCCTCAGGTAAAGCCGGCTAAATACCGAAACAAATACCTATTGTTTGTGTAGACACC 2640
DB PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 929
QY 2641 TTTTCAGGATGGGTAGAGCTTATCTCTAATAAGAAAGACTTCAACCGTGGTGGCTAAA 2700
DB PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 949
QY 2701 AAAATACGCAAGAAATTTTCCAGATTTGGATACCTAAGCTTAAGGTTCAGACAAAT 2760
DB LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 969

QY 2761 GGTCCAGCTTTTGTGCCCCAGGTAAAGTCAGGACTGGCCAGATATTTGGGGATTGATGG 2820
DB GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyLeuAspTrp 989
QY 2821 AAACTGCATTTGTCATACAGACCCCAAGCTCAGACAGGTAGAGAGATCAATAGAACC 2880
DB LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 1009
QY 2881 ATTAAGAGACCCCTTACTAAATTGACCGCGAGACTGGCGTTAATGATTGGATAGCTCTC 2940
DB IleLysGluThrLeuThrLysLeuThrGluThrGlyIleAsnAspTrpIleAlaLeu 1029
QY 2941 CTGCCCCCTTTGCTTTTGTAGGTTAGGAACACCCCTCGACAGTTTGGGCTGACCCCTAT 3000
DB LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1049
QY 3001 GAATTACTCTACCGGGGACCCCCCAATTTGGTAGAAAATGCTTCTGTACATAGTCTGAC 3060
DB GluLeuLeuTyrGlyGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp 1069
QY 3061 GTGCTGCTTTCCAGACCTTTGTTCTTAGGCTCAAGCCTTGTAGTGGGTGAGACAACGA 3120
DB ValLeuLeuSerGlnProLeuPheSerArgLysAlaLeuGluTrpValArgGlnArg 1089
QY 3121 GCGTGGAGGCAACTCCGGAGGCTTACTCAGGAGGAGGACTTTCAGATCCACATCGT 3180
DB AlaTrpLysGlnLeuArgGluAlaTyrSerGlyGluGlyAspLeuGlnValProHisArg 1109
QY 3181 TTCCAAGTGGAGATTCAAGTCTAGTCTAGGACCCACCGTCAGGAAACCTCGACACTCGG 3240
DB PheGlnValGlyAspSerValTyrValArgArgHisAlaGlyAsnLeuGluThrArg 1129
QY 3241 TGAAGGGCCCTTATCTCGTACTTTTCACCAACCAACGGCTGTGAAGTTCGAAGGAATC 3300
DB TrpLysGlyProTyrLeuLeuLeuLeuThrProThrAlaValLysValGluGlyLe 1149
QY 3301 TCCACCTGGATCCATGTCATCCACGTTTAAACCCGCGCCACTCCCGATTTCGGGTGGAAA 3360
DB SerThrTrpIleHisAlaSerHisValLysLeuAlaProProAspSerGlyTrpArg 1169
QY 3361 GCGAAAGACTGAAATCCCTTAAGTTCGCTTCATCGCTGCTCATCGTGGTTCCTTACTGTC 3420
DB AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn 1189
QY 3421 AATAAC 3426
DB AsnAsn 1191
RESULT 11
Q9Q1X3_9GAMR
ID Q9Q1X3_9GAMR PRELIMINARY; PRT; 2376 AA.
AC Q9Q1X3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Type C proviral gag, pol and env genes and LTR (class B, clone 43).
OS Porcine endogenous retrovirus.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Gammaetrovirus; 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Type C;
RX MEDLINE=20219394; PubMed=10756014;
DOI=10.1128/JVI.74.9.4028-4038.2000;
RA Czaderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;
RT "Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells";
RL J. Virol. 74:4028-4038(2000).
DR EMBL; AJ133818; CAB65341.1; -; Genomic_DNA.
DR HSSP; P03385; 1MOF.
DR SMR; Q9Q1X3; 7-98, 2225-2277.

DR MEROPS; A02.020; -
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006310; F:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR GO; GO:0019068; P:viral assembly; IEA.
DR InterPro; IPR002050; Env polypeptide.
DR InterPro; IPR008981; FnuIIYrecept-bind.
DR InterPro; IPR003036; Gag_p30.
DR InterPro; IPR000840; G_retro.matrix.
DR InterPro; IPR001995; Peptidase A2 cat.
DR InterPro; IPR001969; Pept Asp_AS_
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01140; Gag_MA; 1.
DR Pfam; PF02093; Gag_p30; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF00429; TIV_coat; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00879; RNASE H; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.
FT CHAIN 1 524 Gag protein.
FT CHAIN 525 1761 Pol protein.
FT CHAIN 1720 2376 Env protein.
SQ SEQUENCE 2376 AA; 266550 MW; C6D16E368BA49ABF CRC64;

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q9Q1X3_9GAVR (1-2376)

Alignment Scores:
Pred. No.: 0 Length: 2376
Score: 5861.00 Matches: 1108
Percent Similarity: 98.5% Conservative: 17
Best Local Similarity: 97.0% Mismatches: 17
Query Match: 94.8% Indels: 0
DB: 2 Gaps: 0

QY 1 ATGGGTGCCACAGGGCAACACAGTATCCATGGACTACCCGAGAACAGTGTGACTTGGGA 60
DB 574 MetGlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgArgThrValAspLeuAla 593
QY 61 GTGGGACGGGTAAACCCACTCGTTCTGGTCATACCTGAGTCCCGACGACCCCTCTAGGT 120
DB 594 ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly 613
QY 121 AGAGACTATTATGACCAAGATGGGAGACAAATTTCTTTGAAACAAGGMAACCAAGATG 180
DB 614 ArgaspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 633
QY 181 TCTCAAATAACAAACCTATCACTGTGTGTGACCTCCCAATTTAGATGACGAATATCGACTA 240
DB 634 SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 653
QY 241 TACTCTCCCTAGTAAAGCCCTGATCAATATACATTTCTGGTTGGACAGTTTCCCCAA 300
DB 654 TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln 673
QY 301 GCCTGGGCAGAAACCGCAGGAGTGGTTTGGCAAGCAAGTTCCTCCCAAGAGTTATTCAA 360
DB 674 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 693

QY 361 CTGAGGCCAGTGCACACAGTGTCTCAGTCAGAGTACCCCTTGAGTAAAGAGCTCAA 420
DB 694 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg 713
QY 421 GAGGAATTCCGGCGCATGTCCAAAGATTAAATCAACAGGGCATCTAGTTCTCTGTCCAA 480
DB 714 GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyLeuValProValGln 733
QY 481 TCTCCCTGGAAATCTCCCTGCTTACCGGTTAGAAAGCTGGGACTAATAGACTATCGACCA 540
DB 734 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 753
QY 541 GTAAGGACTTGAGAGAGGTCAATAACGGGTGCAGGATATACACCCAAAGTCCCGAAC 600
DB 754 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 773
QY 601 CTTTATAACCTCTTTGTGTCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGACTTA 660
DB 774 ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeu 793
QY 661 AAGGATCCCTTCTCTGCTGAGATTACACCCACTAGCCAAACCACTTTTTCCTTCGAA 720
DB 794 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 813
QY 721 TGGAGAGATCCAGTACGGGAAGAACCGGGCAGCTCACCTGGACCCGACTGCCCCCAAGG 780
DB 814 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 833
QY 781 TTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTCAGG 840
DB 834 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 853
QY 841 ATCCAACACCTCAGGTGACCTCTCTCCAGTACGTGGATGACCTGCTGTGGAAATGTCTGACCTAGGC 900
DB 854 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAla 873
QY 901 ACCAAACAGGACTGCTTAGAGGACCAAGCCACTACTGCTGGAAATGTCTGACCTAGGC 960
DB 874 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 893
QY 961 TACAGAGCTCTGCTAAGAGGCCACAGATTTCAGAGAGAGAGTAAACATCTTGGGGTAC 1020
DB 894 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 913
QY 1021 AGTTTGGGACGGGACGAGATGCTGACGGAGCAGCGAAGAACTGTAGTCCAGATA 1080
DB 914 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 933
QY 1081 CCGGCCCCAACCCACAGCCCAACAAATCAGAGAGTTCCTTGGGGACAGCTGGATTTTGCAGA 1140
DB 934 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 953
QY 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAAACCAAGAAAAA 1200
DB 954 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 973
QY 1201 GGGGAATTCCTCTGGGCTCTGAGCACACAGAAGCATTTTGATGTATCAAAAAAGCCCTG 1260
DB 974 GlyGluPheSerTrpAlaProGluHisGlnLysThrPheAspAlaIleLysLysAlaLeu 993
QY 1261 CTGAGCCGACCTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTTACCTTTATGTGAT 1320
DB 994 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 1013
QY 1321 GAGGCTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGGACCATGAGAGAGACT 1380
DB 1014 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 1033
QY 1381 GTCCCTACCTGTCAAAGAGCTCGATCTGTAGCCAGTGGTTGGCCCATATGCTGTAAG 1440
DB 1034 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTyrProValCysLeuLys 1053

1441 GCTATCGCAGCTGTGGCCATCACTGGTCAAGGACCGTGACAAATTCAGCTTTGGGACAGAA 1500
1054 AlalileAlaValAlaIleLeuValIlysAspAlaAspLysLeuThrLeuGlyGlnAsn 1073
1501 ATAACCTGTAATAGCCCCCATCGATTGGAGAACATCGTTCCGACAGCCCCACAGACCGATGG 1560
1074 IleThrValIleAlaProHisAlaLeuGlnAsnIleValArgGlnProProAspArgTrp 1093
1561 ATGACCAACGCCCGCATGACCCACTATCAAAAGCTGCTTCTCAGAGAGAGGGTCAACGTC 1620
1094 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 1113
1621 GCTCCACAGCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGAACAGTG 1680
1114 AlaProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 1133
1681 ACTCATGATGCCATCAACTATTGATTGAGGAGACTGGGTCGCAAGGACCTTACAGAC 1740
1134 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyAlaArgLysAspLeuThrAsp 1153
1741 ATACCGCTGACTGGAGAAGTCTAACTGGTTCACTGACGGAAGCAGCTATGTGGTGA 1800
1154 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 1173
1801 GGTAAAGGATGGTGGGGCGCGTGGTGGACGGGACCCGACAGACTCTGGGCCAGCAGC 1860
1174 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrHisThrIleTrpAlaSerSer 1193
1861 CTGCGCGAAGAACTTCAGACAAAAGCTGAGCTCATGCGCTCAGCAGCTTTCGGG 1920
1194 LeuProGluGlyThrSerAlaGlnLysAlaGlnLeuMetAlaLeuThrGlnAlaLeuArg 1213
1921 CTGCGCGAAGGAAATCATATAACATTTATACGACAGCAGGTATGCTTTGCGACTGCA 1980
1214 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 1233
1981 CAGTACATGGGGCATCTATAACAAAGGGGTGTCTTACCTCAGCAGGAGGGAATA 2040
1234 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 1253
2041 AAGAACAAAGAGAAATCTAAGCCTATTAGAGCCGTACATTTACCAAAAGGCTAGCT 2100
1254 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 1273
2101 ATTATACACTGTCTGGACATCAAGAAAGCTAATCATATCCAGAGAAACCCAGATG 2160
1274 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 1293
2161 GCTGACCGGTTCGAAGCGGCGCCAGGGTGTAACTTCTGCTTATTAAGAAATG 2220
1294 AlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleIleGluThr 1313
2221 CCCAAAGCCCAAGAACCCAGACGACACTACACCTCAGAACGCTGGCAAGAGATAAAAG 2280
1314 ProLysAlaProGluProArgArgLysGlnThrLeuGluAspTrpGlnGluIleLysLys 1333
2281 ATAGACCACTCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGCGGAAGGAATC 2340
1334 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 1353
2341 CTGCCCCCAAGAGAGGGTGTAGTAATATGTCCAACAGATACATCGTCAACCCACCTAGGA 2400
1354 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 1373
2401 ACTAAACACCTGAGGAGTGTGTCAGAACATCCCTTATCATCTTCTGAGGCTTACCAGGA 2460
1374 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 1393
2461 GTGGCTGACTCGGTGGTCAACAACTGTGTGCCCTGCCAGCTGGTGTATGTAATCTCTTCC 2520
1394 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 1413
2521 AGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGGCGCTCACTGGGAAGTGGAC 2580

1414 ArgValProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 1433
2581 TTCACTCAGGTAAAGCCGCTAAATACGGAACCAATACCTATTGTTTGTGTAGACACC 2640
1434 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 1453
2641 TTTTCAGATGGGTAGAGGCTTATCTTACTAAAGAAAGAGACTTCAACCCGTCGTGCTAAA 2700
1454 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 1473
2701 AAAATACCTGAAGAAATTTTCCAGATTTCGAATATACCTAAGGTAAATAGGTCAGACAA 2760
1474 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 1493
2761 GGTCCAGCTTTGTTCGCCAGGTAAAGTCAGGACTGCGCAAGATATTGGGGATTGATTGG 2820
1494 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 1513
2821 AAATGCTATTGTGCATACAGACACCCAAAGCTCAGGACAGGCTAGAGAGATGAATAGAAC 2880
1514 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnLysThr 1533
2881 ATTAAGAGACCTTACTAAATTGACCCGCGAGAGACTGGCTTAATGATTGATGCTCTC 2940
1534 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu 1553
2941 CTGCGCTTGTGCTTTTAGGTTAGGAACACCCCTGGACAGCTTGGGCTGACCCCTAT 3000
1554 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1573
3001 GAATTACTCTACCGGGGACCCCTTCTGTTAGAAATTCCTCTGACATAGTCTGCTGAC 3060
1574 GluLeuLeuTyrGlyGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp 1593
3061 GTGCTGTTTCCAGCTTTGTTCTTAGGCTCAAGGCACTTGAAGTGGTGAGACAACGA 3120
1594 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1613
3121 GCTGAGGAGCACTCCGGAGGCTACTCAGGAGGAGGACTTGCAGATCCACATCGT 3180
1614 AlaTrpLysGlnLeuArgGluAlaTyrSerGlyGluGlyAspLeuGlnValProHisArg 1633
3181 TTCACAGGAGGAGATTCAGTCTACGTTAGACGCCACCGCTCAGGAAACCTCGAGACTCG 3240
1634 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1653
3241 TGAAGGCGCTTATCTCGTACTTTTGAACCAACCGCTGTGAAGTGAAGGAATC 3300
1654 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1673
3301 TCCACCTGGATCCATCCATCCACCGTTAAACCGCGCCACCTCCGATTCGGGGTGA 3360
1674 SerThrTrpIleHisAlaSerHisValLysLeuAlaProProProAspSerGlyTrpArg 1693
3361 GCCGAAAGACTGAAATCCCTTAAAGCTTCGCTCATCGCTGCTTCTTCTACTCTGTC 3420
1694 AlaGluThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn 1713
3421 AATAAC 3426
1714 AsnAsn 1715
RESULT 12
Q8UM99_9GAMR PRELIMINARY; PRT; 1144 AA.
ID Q8UM99_9GAMR
AC Q8UM99;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Poi protein.
GN Name=poi;
OS Porcine endogenous retrovirus.

OC Viruses; Retrovirus; Retroviridae; Gammatetrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20219394; PubMed=10756014;
RX DOI=10.1128/JVI.74.9.4028-4038.2000;
RA Czauderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;
RT "Establishment and characterization of molecular clones of porcine
RT endogenous retroviruses replicating on human cells.";
RL J. Virol. 74:4028-4038(2000).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21851083; PubMed=11861838;
RX DOI=10.1128/JVI.76.6.2714-2720.2002;
RA Niebert M., Rogel-Gaillard C., Chardon P., Toenjes R.R.;
RT "Characterization of Chromosomally Assigned Replication-Competent Gamma
RT Porcine Endogenous Retroviruses Derived from a Large White Pig and
RT Expression in Human Cells";
RL J. Virol. 76:2714-2720(2002).
DR EMBL; AJ279056; CAC82502.1; -, Genomic_DNA.
DR HSSP; P03355; 116J.
DR MEROPS; A02.020; -.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0004523; P:ribonuclease H activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003964; P:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS00879; RNASE H; 1.
SQ SEQUENCE 1144 AA; 127926 MW; 5F3AA5AEFB98038 CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1144
Score: 5840.50 Matches: 1107
Percent Similarity: 97.8% Conservative: 12
Best Local Similarity: 96.8% Mismatches: 24
Query Match: 94.5% Indels: 1
DB: 2 Gaps: 1

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q8UM99_9GAMR (1-1144)

QY 1 ATGGGTGCGCAGGGCAACACAGTATCCATGGACTACCCGAGAACAGTTTGACTTGGGA 60
DB 1 MetGlyAlaThrGlyGlnGlnGlnTrpTrpThrArgArgThrValAspLeuGly 20

QY 61 GTGGGACGGGTAAACCACTCGTTTCTGTCATACCTGAGTGCACCGACCCCTCTTAGGT 120
DB 21 ValGlyArgValThrGlnSerPheLeuGlyProGluGlyProGluProLeuLeuGly 40

QY 121 AGAGACTTATTGACCAAGATGGGAGCAAAATTTCTTTTGAAACAGGGAACCGAAGTG 180
DB 41 ArgAspLeuLeuThrThrMetGlyArgGlnPhePheGluGlnGlyThrProGluVal 60

QY 181 TCTGCATAATACAACTATCACTGTGTGGACCTCCCAATTAGATGACGAATATCGACTA 240
DB 61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspGlnLysArgLeu 80

QY 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATTTCTGTTGGAAACAGTTTCCCAA 300
DB 81 TyrSerSerLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln 100

QY 301 GCCTGGGAGAAACCGCAGGATGGGTTGGCAAGCAAGTTCCCCCAAGTTATTCAA 360
DB 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120

QY 361 CTGAAGGCCAGTGCCACACCAAGTGTGTCAGTCAGACAGTACCCCTTGAGTAAAGAGCTCAA 420

Db 121 LeuLysAlaSerAlaAlaProValSerValArgGlnTrpProLeuSerLysGluAlaArg 140
QY 421 GAAGGAATTCGGCGCATGTCCTCCAAAGATTAAATCCACAGGCGATCCTAGTTCCTGCCAA 480
Db 141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 160
QY 481 TCTCCCTGGGAATCTCTCCCTGCTACCGGTTAGAAAGCTGGGACTAATGACTATCGACCA 540
Db 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
QY 541 GTACAGGACTTGCAGAGGTCATATAAAGCGGTGCAGGATATACACCAACAGTCCCGAAC 600
Db 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200
QY 601 CCTTATTAACCTCTTGTGTCTCTCCACCCCAACGAGCTGGGTATACAGTATTCGACTTA 660
Db 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu 220
QY 720 661 AAGGATCGCTTCTCTGCTGAGATTACACCCACTAGCAACCACTTTTTCCTTCGAA 720
Db 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
QY 780 721 TGAAGAGATCCAGGTACGGGAAGAACCGGCGAGCTCACCTGGACCCGACTGCCCAAGGG 780
Db 260 241 TrpArgAspProGlyAlaGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
QY 840 781 TTCAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGG 840
Db 280 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280
QY 900 841 ATCCAACACCTCAGGTGACCTCTCCAGTACGTGATGACCTGCTCTCGCGGGAGCC 900
Db 300 281 IleGlnHisProGlnValThrLeuLeuGlnTrpValAspAspLeuLeuAlaGlyAla 300
QY 960 901 ACCAAACAGGACTCTTTAGAAAGGCACGAAGSCACTACTGCTGGAATTTCTGACTAGGC 960
Db 320 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 320
QY 1020 961 TACAGACCTCTGCTAAGAGGCCAGATTTCGAGGAGAGAGGTAAACATCTTGGGTAC 1020
Db 340 321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 340
QY 1080 1021 AGTTTGGGGACGGGACGATGGCTGACGAGGACGCGAAGAAACCTGTAGTCCAGATA 1080
Db 360 341 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysArgThrValValGlnIle 360
QY 1140 1081 CCGGCCCAACCAACAGCCAAACAAATGAGAGAGTTCCTTGGGACAGCTGGATTTTGCAGA 1140
Db 380 361 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 380
QY 1200 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCTTACCCGCTAACCAAGAAAAA 1200
Db 400 381 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 400
QY 1260 1201 GGGGAATTCCTCGGCTCCTGACGACGAGGATTTGATGCTATCAAAAAGSCCCTG 1260
Db 420 401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
QY 1320 1261 CTGACGCGACCTGCTGCGCCCTCCCTGACGTAATAAACCTTTTACCCCTTATGTGAT 1320
Db 440 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440
QY 1380 1321 GAGGTAAAGGAGTAGCCCGGGATTTTAAACCAACCCCTAGACCCATGAGGAGACT 1380
Db 460 441 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 460
QY 1440 1381 GTCCCTACCTGTCAAAGAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGCTGAAG 1440
Db 480 461 ValAlaLysLeuSerLysLysLeuAspProValAlaSerGlyTyrProValCysLeuLys 480
QY 1500 1441 GCTATCGCAGCTGTGGCCATATCTGGTCAAGGACCTGCAAAATGACTTTTGGACAGAT 1500

Db 481 AlalleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500
QY 1501 ATAACGTAAATAGCCCCCATTCATTCGAGAACATCGTTCCGACAGCCCCCAGACGATGG 1560
Db 501 IleThrValIleAlaProHisAlaLeuGlnAsnIleValArgGlnProProAspArgTrp 520
QY 1561 ATGACCAACGCCCGCATGACCACTATCAAGACCTGCTTCACAGAGAGGGTCACGTTTC 1620
Db 521 MetThrAsnAlaArgMetThrHisGlnSerLeuLeuThrGluArgValThrPhe 540
QY 1621 GCTCCACGAGCGCTCTCAACCCCTGCCACTCTCTGCTGCTGAAGAGACTGTAACCCAGTG 1680
Db 541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 1681 ACTCATGATCCCACTCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACTTACAGAC 1740
Db 561 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580
QY 1741 ATACCGCTGACTGGAGAAGTCTTAACCTGGTTCATGACGGAAGAGCTATGTGGTGAA 1800
Db 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrrValValLys 600
QY 1801 GGTAAAGAGGATGGCTGGGGCGGCTGTGTGGACGGGCCCCACGATCTGGCCAGCAGC 1860
Db 601 GlyLysArgMetAlaGlyProProValValAspGlyThrArgThrIleTrpAlaSerSer 620
QY 1861 CTGCGGGAAGGAATTCAGCACAAAGGCTCATGCTCATGGCCCTCAGCAAGCTTTCGGG 1920
Db 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 1921 CTGCGGGAAGGAATTCATAAATTTATACGACAGCAGGTATGCTTTCGAGCTGCA 1980
Db 641 LeuAlaGluGlyLysSerIleAsnIleTyrrThrAspSerArgTyrrAlaPheAlaThrAla 660
QY 1981 CACGTACATGGGCATCTAATAACAAGGGGTGCTTACCTCAGCAGGGAGGGAATA 2040
Db 661 HisValHisGlyAlaIleTyrrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 2041 AAGAACAAAGAGGAATTTCTAAGCTATTAGAAGCGGTACATTTACAAAAGGCTAGCT 2100
Db 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 700
QY 2101 ATTATACATCTCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAACACGATG 2160
Db 701 IleIleHisCysProGlyHisGlnLysAlaLysAspProIleSerArgGlyAsnGlnMet 720
QY 2161 GCTGACCGGTTGCAAGCAGGACCCCGGGTGTAACTTCTGCTTATATAAGAAATG 2220
Db 721 AlaAspArgValAlaLysGlnAlaGlnGlyValAsnLeuLeuProMetIleGluThr 740
QY 2221 CCCAAAGCCCCAGAACCCAGACGACGTACACCCCTAGAAGACTGGCAAGAGATAAAAAG 2280
Db 741 ProLysAlaProGluProGlyArgGlnTyrrThrLeuGluAspTrpGlnGluIleLysLys 760
QY 2281 ATAGACAGATTCTCGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGGAATC 2340
Db 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrrThrSerAspGlyLysGluIle 780
QY 2341 CTGCCCCACAAAGAGGGTTAGATATGTCCACAGATACATGCTTAACCCACCTAGGA 2400
Db 781 LeuProHisLysGluGlyLeuGluTyrrValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 2401 ACTAAACACCTGCAGCAGTGGTCAGAAACATCCCTTATCATGTTCTGAGCGTACCAGGA 2460
Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrrHisValLeuArgLeuProGly 820
QY 2461 GTGCGTCACTCGGTGGTCAAAACATTTGTGTGCTGCCAGCTGGTTTAACTGCTTAATCTCTCC 2520
Db 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY 2521 AGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCACTGGGAAGTGAC 2580
Db 841 ArgIleProProGlyLysArgLeuArgLysSerHisProGlyAlaHisTrpGluValAsp 860

QY 2581 TTCACTGAGGTAAAGCCGGCTAAATACGAAACAAATACCTATTGGTTTGTGTAGACACC 2640
Db 861 PheThrGluValLysProAlaLysTyrrGlyAsnLysTyrrLeuLeuValPheValAspThr 880
QY 2641 TTTTCAGAGTGGGTAGAGCTTATCTTACTTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 2700
Db 881 PheSerGlyTrpValGluAlaTyrrProThrLysLysGluThrSerThrValValAlaLys 900
QY 2701 AAAATACCTGGAAGAAATTTTCCAAAGATTGGAATACCTAAGTAATAGGTCTGACACAT 2760
Db 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
QY 2761 GGTCCAGCTTTTGTGTCGCCAGGTAAAGTACAGGACTGCGCAAGATATTGGGATTCATTGG 2820
Db 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY 2821 AAACGTGATTGTGATACAGACCCCAAGCTCAGAGCAGGTAGAGAGATGAATAGAAC 2880
Db 941 LysLeuHisCysAlaTyrrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY 2881 ATTAAGAGACCCCTTACTTAAATTGACCGCGAGACTGGCGTTAATGATTGGATGACTCTC 2940
Db 961 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu 980
QY 2941 CTGCGCTTTGCTGCTTTTAGGGTTAGGAACAACCCCTCGACAGTTTGGCTGACCCCTAT 3000
Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrr 1000
QY 3001 GAATTACTCTACGGGGGACCCCGCCCATTTGGTAGAAATGCTCTGTACATAGTCTGAC 3060
Db 1001 GluLeuLeuTyrrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSerAlaAsp 1020
QY 3061 GTGCTGCTTTCCAGACCTTTGTTCTTAGGCTCAAGGCACTTGAGTGGGTGAGCAACGA 3120
Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY 3121 GCGTGGAGGCAACTCCGGGAGGCGCTACTCAGGAGGAGAGACTTGCAGATCCACATCGT 3180
Db 1041 AlaTrpLysGlnLeuArgGluAlaTyrrSer---GlyGlyAspLeuGlnValProHisArg 1059
QY 3181 TTCAGTGGAGATTCAGTCTAGTTCAGCCACCGCTGAGGAAACCTCGAGACTCGG 3240
Db 1060 PheGlnValGlyAspSerValTyrrValArgHisArgAlaGlyAsnLeuGluThrArg 1079
QY 3241 TGGNAGGGCCCTTATCTCGTACTTTTCACACACCAACGCGCTGGAAGTCAAGGAATC 3300
Db 1080 TrpLysGlyProTyrrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1099
QY 3301 TCCACCTGGATCCATGCATCCACGTTTAAACCGCGCCACTCCCGATTCGGGGTGGAAA 3360
Db 1100 ProThrTrpIleHisAlaPheHisValLysProAlaProProSerAspSerGlyTrpLys 1119
QY 3361 GCCGAAAGACTGAAATCCCTTAAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTC 3420
Db 1120 AlaGlnLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrrSerVal 1139
QY 3421 AATAACCTCTCA 3432
Db 1140 AsnAsnSerSer 1143

RESULT 13

Q8UM96_9GAMR
ID Q8UM96_9GAMR PRELIMINARY; PRT; 1147 AA.
AC Q8UM96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol protein.
GN Name-pol;
OS Porcine endogenous retrovirus.
OC Viruses; Retrovirdae; Gammaretrovirus;
OC 1-Mammalian type C virus group.

OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20219394; PubMed=10756014;
RA DOI=10.1128/JVI.74.9.4028-4038.2000;
RA Czauderna F., Fischer N., Boller K., Kurth R., Toenjes R.R.;
RT "Establishment and characterization of molecular clones of porcine
RT endogenous retroviruses replicating on human cells.";
RL J. Virol. 74:4028-4038(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21851083; PubMed=11861838;
RX DOI=10.1128/JVI.76.6.2714-2720.2002;
RA Niebert M., Rogel-Gaillard C., Chardon P., Toenjes R.R.;
RT "Characterization of Chromosomally Assigned Replication-Competent Gamma
RT Porcine Endogenous Retroviruses Derived from a Large White Pig and
RT Expression in Human Cells.";
RL J. Virol. 76:2714-2720(2002).
DR EMBL: AJ279057; CAC82505.2; -, Genomic_DNA.
DR MEROPS: A02.020; --
DR GO: GO:004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0006310; P:DNA recombination; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR001995; Peptidase_A2_cat.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR004477; RVTse.
DR Pfam: PF00075; RNaseH; 1.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00077; RVP; 1.
DR Pfam: PF00078; RVT; 1.
DR PROSITE: PS0175; ISP_PROT_RETROV; 1.
DR PROSITE: PS0879; RNASE_H; 1.
SQ SEQUENCE 1147 AA; 128404 MW; C6374B0CE460901D CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1147
Score: 5831.00 Matches: 1104
Percent Similarity: 98.4% Conservative: 20
Best Local Similarity: 96.7% Mismatches: 18
Query Match: 94.3% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q8UM96_9GAMR (1-1147)

Qy 1 ATGGGTGCCACAGGGCAACAACAGTATCCATGGACTACCCGGAACAGTTGACTTGGGA 60
Db 1 MetGlyAlaThrGlyGlnArgGlnProTrpThrArgArgThrValAspLeuVal 20

Qy 61 GTGGGAGGGTAACCCACTCGTTCTGCTGCTATACCTGAGTCCCGAGCACCCCTCTTAGT 120
Db 21 ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly 40

Qy 121 AGAGACTTATTGACCAAGATGGGAGCAAAATTTCTTTTGAACAGGGAACCCAGAAAGTG 180
Db 41 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 60

Qy 181 TCTGCAATATACAAACCTATCACTGTGTGTGACCTCCATAGATGAGCAATATCGACTA 240
Db 61 SerValAspAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 80

Qy 241 TACTCTCCCTAGTAAGCCCTGATCAAAATATCAATTTCTGTTGGACAGTTCCTCCCAA 300
Db 81 TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGlnPheProGln 100

Qy 301 GCCTGGGAGAAACCGAGGGATGGTTTGGCAAGCAAGTATCCCCCAACAAGTTATTCAA 360
Db 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120

Qy 361 CTGAGGCCAGTGCCACACAGTGTCTAGTCAGACAGTACCCCTTGAGTAAAGAAGCTCAA 420
Db 121 LeuLysAlaSerAlaThrProValSerValArgGlnTrpProLeuSerArgGluAlaArg 140

Qy 421 GAAGGAATTCGGCGCATGTCCTCAAGATTAATCAACAGGGCATCTCTAGTCTCTGTCCAA 480
Db 141 GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyLeuValProValGln 160

Qy 481 TCTCCCTGGGAATACTCCCTGCTACCGGTTAGAAGCCTGGGACTAATGACTATCGACCA 540
Db 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTrpArgPro 180

Qy 541 GTACAGGACTTGGAGAGGGTCAATAAAGCGGTGAGGATATACACCAACAGTCCCCAAC 600
Db 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200

Qy 601 CCTTATAACCTCTTGTGTCTCTCCCAACCCGAGCTGGTATACAGTATTCGACTTA 660
Db 201 ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTrpThrValLeuAspLeu 220

Qy 661 AAGATGCTCTTCTGCTGAGATTACACCCACTAGCCCAACCACTTTTTCCTTCGAA 720
Db 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240

Qy 721 TGGAGAGATCCAGTACGGGAAGAACCGGCGAGCTCACCTGGACCCGACTGCCCCCAAGGG 780
Db 241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260

Qy 781 TTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGACAGACTGGCCCACTTCAGG 840
Db 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280

Qy 841 ATCCAACACCTCAGGTGACCTCTCTCCAGTACGTGATGACCTGCTTCTGGCGGAGCC 900
Db 281 IleGlnHisProGlnValThrLeuLeuGlnTrpValAspAspLeuLeuLeuAlaGlyAla 300

Qy 901 ACCAAACAGGACTCTGTAGAGCGACCAAGGCACACTACTGCTGGAATTTGTCTGACTAGGC 960
Db 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuSerAspLeuGly 320

Qy 961 TACAGAGCCTCTGTGAAGAAGGCCACAGATTTCGAGGAGAGAGGTAACATCTTGGGTAC 1020
Db 321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTrpLeuGlyTyr 340

Qy 1021 AGTTTGGGAGCGGCGAGCGATGCTCAGCGAGCGACCGAGAGAACTGTAGTCCAGATA 1080
Db 341 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 360

Qy 1081 CCGGCCCAACACAGCCCAACAAATGAGAGAGTTTTTTGGGGACAGCTGGATTTCGAGA 1140
Db 361 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 380

Qy 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCTCTCTACCCGCTAACCAAGAAAA 1200
Db 381 LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLys 400

Qy 1201 GGGGAATTCCTCGGCTCTGAGCACACAGAGGCATTTGATGCTATCAAAAGGCCCTG 1260
Db 401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420

Qy 1261 CTGAGCCCATCTGCTGGCCCTCCCTGACCTAAACCCCTTTTACCTTTTATGTGGAT 1320
Db 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440

Qy 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCCCTAGGACCATGAGAGACCT 1380
Db 441 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 460

Qy 1381 GTCGCCCTACTGTCAAAGAAGCTCGATCCTGTACCGAGTGGTTGGCCCATATGCTGAG 1440
Db 461 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLys 480

QY 1441 GCTATCGAGCTGTGGCCATCTACTGGTCAAGGACCGCTGACAAATGACTTTGGGACAGAAAT 1500
Db AlaileAlaAlaValAlaileLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAen 500
QY 1501 ATAACGTAAATAGCCCCCATGCTGAGAGAACATCGTTGCGGACGCCCCACAGACCGATGG 1560
Db IleThrValIleAlaProHisAlaLeuGlnAenGlnAenIleValArgGlnProProAspArgTrp 520
QY 1561 ATGACCAACGCCCGCATGACCCACTATCAAGACCTGCTTCTCAGACAGAGAGGCTCACGCTTC 1620
Db MetThrAenAlaArgMetThrHisTyGlnSerLeuLeuLeuThrGluArgValThrPhe 540
QY 1621 GCTCACAGCGGCTCTCAACCCCTGCCACTCTTCTGCTGAGAGAGACTGATGAACCAAGTG 1680
Db AlaProProAlaAlaLeuAenProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 1681 ACTCATGATTCGCCATCAACTATTGATTGAGGAGACTGGGTCGCCAGGACCTTACAGAC 1740
Db ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580
QY 1741 ATACCGCTGACTGGAGAAGTGCTAACTGTTCTACTGACGGAAGCAGCATGTGTGGTGAA 1800
Db IleProLeuThrGlyGluValIleuThrTrpPheThrAspGlySerSerTyValValGlu 600
QY 1801 GGTAAAGAGATGCTGGGGCGCGGTGTGTGACCGGACCCGCCACGATCTGGGCCAGCAGC 1860
Db GlyLysArgMetAlaGlyAlaAlaValAlaAspGlyThrArgThrIle**AlaSerSer 620
QY 1861 CTGCGGAGAGAACTTCAGACAAAGGCTGAGCTCATGCGCCCTCAGCGAAGCTTTCGGG 1920
Db LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 1921 CTGCGCGAAGGAAATCATATAAATTTATACGACAGCAGGATGCTTTCGACCTGCA 1980
Db LeuAlaGluGlyLysSerIleAenIleTyThrAspSerArgTyAlaPheAlaThrAla 660
QY 1981 CACGTACATGGGCGCATCTATAAACAAGGGGTGTCTTACTCAGCAGGAGGAGAAATA 2040
Db HisValHisGlyAlaIleTyLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 2041 AAGAACAAAGAGAAATCTTAAGCCTATTAGACCGCTGACATTTACCAAAAGGCTAGCT 2100
Db LysAenLysAspGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 700
QY 2101 ATTATACACTGCTCTGACACTCAGAAAGCTAAAGATCTCATATCCAGAGGAAACCAAGTG 2160
Db IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAenGlnMet 720
QY 2161 GCTGACCGGGTTCGCAAGCGGACGCCAGGGTGTAACTTCTGCTTATAATAGAAATG 2220
Db AlaAspArgIleAlaLysGlnAlaAlaGlnAlaValAenLeuLeuProIleIleGluThr 740
QY 2221 CCCAAAGCCCCAGAACCCAGACGACACTACACCTAGAACCTGGCAAGAGATATAAAG 2280
Db ProLysAlaProGluProArgArgGlnTyThrLeuGluAspTrpGlnGluIleLysLys 760
QY 2281 ATAGACAGGTTCTCTGAGACTCCGGAAGGACCTGCTTATACCTCAGATGGGAAGGAATC 2340
Db IleAspGlnPheSerGluThrProGluGlyThrCysTyThrSerTyThrSerTyGlyLysGluIle 780
QY 2341 CTGCCCCCAAGAGAGGTTAGATATGTCCAAAGATACATGCTTAAACCCACCTAGGA 2400
Db LeuProHisLysGluGlyLeuGluTyValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 2401 ACTAAACACCTGACGAGGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCTTACAGGA 2460
Db ThrLysHisLeu**GlnLeuValArgThrSerProTyThrHisValLeuArgLeuProGly 820
QY 2461 GTGGCTGACTCGGTGTGCAACACTTGTGTGCCCTGCCAGCTGGTGTAAATGTAATCTCTTCC 2520
Db ValAlaAspSerValValLysHisCysValProCysGlnLeuValAenAlaAenProSer 840
QY 2521 AGAATGCTCCAGGGAAGAGACTAAGGGGAGCCACCCAGGCGCTCACTGGGAAGTGGAC 2580

Db 841 LysValProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
QY 2581 TTCACTGAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGGTTTGTGTAGACACC 2640
Db PheThrLysValLysProAlaLysTyGlyAenLysTyLeuLeuValPheValAspThr 880
QY 2641 TTTTCAGGATGGGTAGAGGCTTATCTTACTAAGAAAGAGACTTCAACCGTGGTGTGCTAAA 2700
Db PheSerGlyTrpValGluAlaTyProThrLysLysGluThrSerThrValValAlaLys 900
QY 2701 AAAATACCTGGAAGAAATTTTCCAAAGATTGGAAATACCTAAGGTAAATAGGTTCAGACAA 2760
Db LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAen 920
QY 2761 GGTCCAGCTTTGTGTGCCAGGTAAAGTCAAGGACTGCGCCAGATATTGGGGATTGATTGG 2820
Db GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY 2821 AAACGTCAATTGTGCATACAGACCCCAAGCTCAGACAGGCTAGAGAGATGAATAGAAC 2880
Db LysLeuHisCysAlaTyArgProGlnSerSerGlyGlnValGluArgMetAenArgThr 960
QY 2881 ATTAAGAGACCCCTTACTAAATTGACCCGCGAGAGACTGGCGTTAATGATTGGATAGCTCTC 2940
Db IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAenAspTrpIleAlaLeu 980
QY 2941 CTGCGCTTGTGCTTTTAGGCTTAGGAACACCCCTGAGCAGTGTGGCTGACCCCTTAT 3000
Db LeuProPheValLeuPheArgValArgAenThrProGlyGlnPheGlyLeuThrProTy 1000
QY 3001 GAATTACTCTACGGGGGACCCCGCTTGGTAGAAATTTGCTCTGACATAGTGTGCTGAC 3060
Db GluLeuLeuTyGlyGlyProProProLeuValGluIleAlaProValHisSerAlaAsp 1020
QY 3061 GTGCTGCTTCCAGCGCTTTGTTCTCTAGCTCAAGGCACCTTGAGTGGGTGAGACAACGA 3120
Db ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY 3121 GCGTGGAGGCAACTCCGGGAGGCTACTCAGGAGGAGAGACTTGCAGATCCCACTCGT 3180
Db AlaTrpLysGlnLeuArgGluAlaTySerGlyGluGlyAspLeuGlnValProHisArg 1060
QY 3181 TTCCAAGTGGGAGATTCAGTCTACGTTAGACGCCCGCTCAGGAAACCTCGAGACTCGG 3240
Db PheGlnValGlyAspSerValTyValArgArgHisArgAlaGlyAenLeuGluThrArg 1080
QY 3241 TGAAGGCGCTTATCTCGTACTTTTGACCCACACCGCTGTGAAAGTCGAAGGAATC 3300
Db TrpLysGlyProTyLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
QY 3301 TCCACCTGGATCCATGCTCCACGCTTAACCGCGCCGCTCCGATTCGGGTGGA 3360
Db SerThrTrpIleHisAlaSerHisValLysLeuAlaProProProAspSerGlyTrpArg 1120
QY 3361 GCCGAAAGACTGAAATCCCTTAAAGTTCGCTCCATCGCTGTTCTTACTCTGTC 3420
Db AlaGluLysThrGluAenProLeuLysLeuArgLeuHisArgLeuValProTySerAen 1140
QY 3421 AATAAC 3426
Db 1141 AsnAen 1142
RESULT 14
Q8Q6U4_9GAMR PRELIMINARY; PRT; 1142 AA.
ID Q8Q6U4_9GAMR
AC Q8Q6U4;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Poi protein.
GN Name:pol;
OS Porcine endogenous retrovirus.

OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
ON NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21851083; PubMed=11861838;
RX DOI=10.1128/JVI.76.6.2714-2720.2002;
RA Niebert M., Rogel-Gallard C., Chardon P., Toenjes R.R.;
RT "Characterization of Chromosomally Assigned Replication-Competent Gamma
RT Porcine Endogenous Retroviruses Derived from a Large White Pig and
RT Expression in Human Cells";
RL J. Virol. 76:2714-2720(2002).
DR EMBL; AF435967; AAL67856.1; -; Genomic_DNA.
DR HSSP; P03355; 1MWL.
DR MEROPS; A02.020; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; F:DNA recombination; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50879; RNase H; 1.
SQ SEQUENCE 1142 AA; 127712 MW; B0F55CA742AAC98F CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1142
Score: 5466.50 Matches: 1042
Percent Similarity: 93.4% Conservative: 26
Best Local Similarity: 91.1% Mismatches: 73
Query Match: 88.4% Indels: 3
DB: 2 Gaps: 2

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q086U4_9GMR (1-1142)

QY 1 ATGGGTGCCACAGGGCAACACAGTATCCATGGATACCCGAGAACAGTGTGGGA 60
Db 1 MetGlyAlaThrGlyGlnGlnThrProThrArgArgThrValAspLeuGly 20

QY 61 GTGGACGGGTAAACCACTGCTTCTGTCTATACCTAGTGCACAGCACCCCTCTAGGT 120
Db 21 ValGlyArgValThrGlnSerPheLeuGlyIleProGluCysProGluProLeuLeuGly 40

QY 121 AGAGACTTATTGACCAAGATGGAGCAAAATTTCTTTGAACAGGGAACCAAGAGTG 180
Db 41 ArgAspLeuLeuThrThrMetGlyArgGlnIlePhePheGluGlnGlyThrProGluVal 60

QY 181 TCTGCAATAACAAACCTATCACTGTGTGACCTCCCAATTAGATGACGAATATCGACTA 240
Db 61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluThrArgLeu 80

QY 241 TACTCTCCCTAGTAAACCTGATCAAAATATACAAATTCCTGGTGGAAACAGTTCCCCAA 300
Db 81 PheSerProProValLysLeuAspGlnAsnIleGlnPheGlySerThrGlnPheProGln 100

QY 301 GCCTGGGAGAAACCGCAGGATGGTTGGCAAGCAAGTTCCCCCAAGTTATTCAA 360
Db 101 AlaLeuAlaGluProAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120

QY 361 CTGAAGCGCAGTCCACCACTGTCTAGTCAGACAGTACCCCTTGTAGTAAAGAAGCTCAA 420
Db 121 LeuLysProSerLeuAlaProValProValArgGlnSerProPheSerLysGluAlaArg 140

QY 421 GAAGAAATTCGGCCCGCATGTCCAAAGATTAATCCAGAGGGATCTAGTTCCTGTCCAA 480
Db 141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleIleValProValGln 160

QY 481 TCTCCCTGGAATACTCCCTGCTACCGGTTAGAAGCCCTGGGACTAATGACTATCGACCA 540

Db 161 SerProThrAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
QY 541 GTACAGGACTTGAGAGAGGTCAATAAAGCGGTGAGGATATACACCCCAAGTCCCGAAC 600
Db 181 ValGlnAspPheGluArgGlyGlnLysArgValGlnAspIleHisProThrValProAsn 200
QY 601 CCTTATAACCTCTTGTCTGCTCTCCACCCCAACGAGCTGTATACAGTATTGGACTTA 660
Db 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTyrThrValLeuAspLeu 220
QY 661 AAGGATCCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAA 720
Db 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
QY 721 TGGAGATCCAGTACGGGAGAAACCGGGCAGCTACCTGGACCCGACTGCCCAAGGG 780
Db 241 TrpArgAspProGlyAlaGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
QY 781 TTCAGAAGCTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTCAGG 840
Db 261 PheLysAsnPheProThrIlePheAspGlnAlaLeuHisArgAspLeuAlaAsnPheArg 280
QY 841 ATCCACACACCTCAGGTGACCTCTCTCCAGTACCTGGATGATGCTTCTTCTGGCGGAGCC 900
Db 281 IleGlnHisProGlnValThrLeuLeuGlnThrValAspAspLeuValLeuAlaAla 300
QY 901 ACCAAACAGGACTCTTAGAAGGACCAAGCAGTACTCTGCGAATGTCTGACTAGGC 960
Db 301 ThrLysGlnAspCysLeuGlnArgProLysGlyLeuLeuValGluLeuSerAspLeuGly 320
QY 961 TACAGAGCTCTGCTAAGAAGGCCAGATTTGACGAGAGAGAGTAAACATCTCTGGGTAC 1020
Db 321 AspArgAlaPheGlyTyrLysAlaHisIleCysProThrGluValThrThrLeuGlyTyr 340
QY 1021 AGTTTGGCGGACGGGACGATGCTGACGAGGACCGAAGAAACCTGTAGTCCAGATA 1080
Db 341 ArgLeuArgGlyArgHisArgTyrLeuThrGluAlaProGlnThrThrValValGlnIle 360
QY 1081 CCGGCCCCAACCCACAGCCAAACAAATGACAGAGAGTTTGGGAGAGCTGGATTTCGAGA 1140
Db 361 ProGlyProThrProAlaLysGlnValArgGluPheLeuGlyThrValIlePheCysArg 380
QY 1141 CTGTGGATCCCGGGTTTGGACCTTAGACCCCACTCTACCCGCTAACCAAGAAAAA 1200
Db 381 LeuThrIleProGlyPheAlaThrLeuProAlaProLeuThrProLeuProLysGluLys 400
QY 1201 GGGGAATTTCTCTGGGCTCTCTGACCAAGGCAATTTGATGCTATCAAAAAGCCCTG 1260
Db 401 GlyGluPheSerTrpAlaLeuGlnHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
QY 1261 CTGAGCGGACCTGCTGCGCCCTCCCTGACGTAACCTTAACCCCTTTACCTTTATGTGAT 1320
Db 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLys-----ThrLeuTyrValAsp 438
QY 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCAACCCCTAGGACCATGAGAGAGACCT 1380
Db 438 LeuSerAlaProAlaLeuAlaLeuProAspValThrLys-----ThrLeuTyrValAsp 438
QY 438 GluArgLysGlyValAlaAlaArgGlyValLeuThrGlnThrLeuGlyProThrArgArgPro 458
QY 1381 GTCGCTACCTGTCTCAAGAAGCTGATCCTCTAGCCAGTGGTGGCCCATATGCTGTAAG 1440
Db 458 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTyrProIleCysLeuLys 478
QY 1441 GCTATCGAGCTGTGGCCATACCTGGTCAAGAGCCTGACCAATTCAGCTTTGGGACAGAT 1500
Db 479 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 498
QY 1501 ATAACTGTAAATAGCCCCCATGCTGAGAAACATCGTTCCGCGACGCCCCAGACCGATGG 1560
Db 499 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTyr 518
QY 1561 ATGACCAACGCCCGCATGACCCACTATCAAAAGCCTGCTTCTCACAGAGGGGTACAGTTC 1620

Db 519 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuLeuThrGluArgValThrPhe 538
QY 1621 GCTCCACGAGCGCTCTCAACCCCTGCCACTCTCTTCCTGCTGAGAGACTGATCAACCACTG 1680
Db 539 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluLeuThrAspGluProLeu 558
QY 1681 ACTCATGATTGCCATCACTATTGATTGAGGAGACTGGGTCCCGAAGGACTTACAGAC 1740
Db 559 ThrHisAspCysHisGlnLeuLeuLeuGluThrGlyValArgLysAspLeuThrAsp 578
QY 1741 ATACCGCTGACTGGAGAGAGTCTAACCTGGTTCCTGACGACGAGCAGCTATGTGTGGAA 1800
Db 579 IleProLeuThrGlyGluProValThrTrpPheThrAspGlySerSerTrpLeuValGlu 598
QY 1801 GGTAAAGAGATGGCTGGCGCGCGTGTGGACGGGACCCGACGATCTGGCGCCAGCAGC 1860
Db 599 GlyAsnLysMetAlaGlyAlaAlaValValAspArgThrProThrIleTrpGlyThrAsn 618
QY 1861 CTGGCGGAAGGAACTTCAGCACAAAGGCTGAGCTCATGCGCCCTCACGCCAAGCTTTCGG 1920
Db 619 LeuProGluArgThrSerSerGlnLysGlyGluLeuIleGlyLeuMetGlnAlaPheArg 638
QY 1921 CTGGCCGAAGGAAATCCATAAACAATTTATACGGACAGCAGTATGCTTTCGCACTGCA 1980
Db 639 LeuGlyGlnGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 658
QY 1981 CACGTACATGGGGCCATCTATAACAAGGGGGTGTCTTACCTCAGCAGGAGGAGGAAATA 2040
Db 659 HisValHisGlyAlaIleTyrThrGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 678
QY 2041 AAGAACAAAGAGGAAATCTTAAGCTATTAGAGCGGTACATTTACCAAAAAGGTAGCT 2100
Db 679 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 698
QY 2101 ATTATACATCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATG 2160
Db 699 IleIleHisCysProGlyHisGlnLysAlaLysAspProIleSerArgLysAsnGlnMet 718
QY 2161 GCTGACCGGGTGGCAAGCAGCGACCCAGGGTGTAACTTCTGCTCTATATAAGAAATG 2220
Db 719 AlaAspArgValAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProMetIleGluThr 738
QY 2221 CCCAAAGCCCAAGAACCCAGACGACGATCACCTTAGAAGACTGGCAGAGATAAAAAG 2280
Db 739 ProLysAlaProGluProGlyArgGlnTyrThrLeuGluAspTrpGlnIleLysLys 758
QY 2281 ATACACAGATTCTCTGAGACTCCGAGAGGACCTGCTATACCTCAGATGGGAAGAAATC 2340
Db 759 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 778
QY 2341 CTGCCCCACAAAGAGGTTAGAAATATGTCCAACAGATATATCTTAACCCACCTAGGA 2400
Db 779 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 798
QY 2401 ACTAAACCTGACGAGTGGTCAGAACATCCCTCTATCATGTTCTGAGGCTACCAAGGA 2460
Db 799 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 818
QY 2461 GTGGCTCAGCTCGGTTCAACATTGTGTCCTCCGACGCTGGTGAATGCTTAATCCTTCC 2520
Db 819 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 838
QY 2521 AGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCACTGGGAAGTGCAC 2580
Db 839 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 858
QY 2581 TTCACTGAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTTGTGTAGACACC 2640
Db 859 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 878
QY 2641 TTTTTCAGATGGGTAGAGGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAA 2700
Db 879 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 898

QY 2701 AAAATACTGGAAGAAATTTTTTCCAGATTTTGAATACCTTAAGGTAAATAGGTCAGACAAT 2760
Db 899 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 918
QY 2761 GGTCCAGCTTTTGTGGCCAGGTAAGTCAGGAGACTGGCCAAAGATATTGGGATGATTGG 2820
Db 919 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 938
QY 2821 AAACCTGCATTGTGCATACAGACCCAAAGCTCAGACAGGTAGAGAGGATCAATAGAAC 2880
Db 939 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 958
QY 2881 ATTAAGAGACCCCTTACTAAATTCACCGCGAGACTGGCGTTAATGATTGGATAGCTCTC 2940
Db 959 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeu 978
QY 2941 CTGCCCTTTGTGCTTTTGTAGGTAGGAACAACCCCTGGACAGTTTGGGCTGACCCCTTAT 3000
Db 979 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 998
QY 3001 GAATTACTCTACGGGGGACCCCGCCCATTTGGTAGAAATTTGCTTCTGTACATAGTGTGAC 3060
Db 999 GluLeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSerAlaAsp 1018
QY 3061 GTGCTGCTTTCCCAAGCCTTTTCTCTAGGCTCAAGGCACCTTGAGTGGGTGAGACAACGA 3120
Db 1019 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1038
QY 3121 GCGTGGAGGCAACTCCGGGAGGCTTACTCAGGAGGAGAGACTTTCAGATCCCACTCGT 3180
Db 1039 AlaTrpLysGlnLeuArgGluAlaTyrSer---GlyGlyAspLeuGlnValProHisArg 1057
QY 3181 TTCCAAGTGGGAGATTCACTGCTAGTACGCCACCGTCAGGAAACCTCGAGACTCGG 3240
Db 1058 PheGlnValGlyAspSerValTyrValArgHisArgAlaGlyAsnLeuGluThrArg 1077
QY 3241 TGAAGGCGCTTATCTCTGCTACTTTTCCACACACCAACCGGTGTGAAGTCAAGAGGAATC 3300
Db 1078 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1097
QY 3301 TCCACCTGGATTCATGATCCACGTTAAACCGCGCCACCTCCCGATTGGGGTGGAAA 3360
Db 1098 ProThrTrpIleHisAlaPheHisValLysProAlaProProSerAspSerGlyTrpLys 1117
QY 3361 GCCGMAAGACTGAAATCCCTTAAGTTCGCTCCATCGCTGGTTCCTTACTCTGTC 3420
Db 1118 AlaGlnLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1137
QY 3421 AATAAACCTCTCA 3432
Db 1138 AsnAsnSerSer 1141

RESULT 15

Q8Q6U7_9GAMR
ID Q8Q6U7_9GAMR PRELIMINARY; PRT; 1139 AA.
AC Q8Q6U7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein.
GN Names; pol;
OS Porcine endogenous retrovirus.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21851083; PubMed=11861838;
RX DOI=10.1128/JVI.76.6.2714-2720.2002;
RA Niebert M., Rogel-gallard C., Chardon P., Toenjes R.R.;
RT "Characterization of Chromosomally Assigned Replication-Competent Gamma
Porcine Endogenous Retroviruses Derived from a Large White Pig and

RT Expression in Human Cells";
 RL J. Virol. 76:2714-2720(2002).
 DR EMBL; AF435966; AAL87853.1; -, Genomic_DNA.
 DR HSSP; P03355; IMML.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006313; P:DNA transposition; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00078; RVT; 1; 1.
 DR PROSITE; PS50879; RNase H; 1.
 SO SEQUENCE 1139 AA; 127325 MW; 3B63C42301C04FAC CRC64;

Alignment Scores:
 Pred. No.: 0 Length: 1139
 Score: 5383.50 Matches: 1054
 Percent Similarity: 94.0% Conservative: 19
 Best Local Similarity: 92.3% Mismatches: 52
 Query Match: 87.1% Indels: 17
 DB: 2 Gaps: 13

US-10-723-552-3_COPY_2307_5741 (1-3435) x Q8Q6U7_9GAMR (1-1139)

QY 41 GAAGAAAGTTGACTTGGGAGTGGACGGGTAAACCACTCTGTTCTGTGATACCTG--- 97
 DB 4 GluAspGlnLeuThrTrpGluLeuAspGlyGluProThrGlyPheTrpValIleLeuLys 23
 QY 98 AGTGCCAGACCCCTCTTAGTAGAGACTTATTGACCAAGATGGAGCACAATTTTC-- 155
 DB 24 Cys-ProSerThrLeuLeuArgValGlnThrTyRAspGlnAspGlyLysProLysPheLe 43
 QY 156 -TTTTGAACAGGGAACCC---AGAAAGTGTCTGCAAAATAACAAACCTATCTGTGTGA 211
 DB 43 uPheGluProArgGluThrGlnLysCysLeu-GlnAsnThrAsnProIleThrValLeuT 63
 QY 212 CCCTCCAAATAGATGACCAATATCGACTATCTCTCCCTAGTAAAGCCTGATCAAAATA 271
 DB 63 hrLeuGlnLeu---AspGluTyRArgLeuTyRSerProLeuValLysProAspGlnAsnI 82
 QY 272 TACAATCTGTTGGAACAGTTTCCCAAGCTCGGCAGAAACCGAGGATGGTTTGG 331
 DB 82 leGlnPheTrpGluGlnPheProLysAlaTrpAlaGluThrAlaGlyMetGlyLeuA 102
 QY 332 CAAGCAAGTTTCCCCCAACAAGTTATTCAACTGAAGCCAGTCCACACACCACTGTCAGTCA 391
 DB 102 laLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaAlaProValSerValA 122
 QY 392 GACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCCGATGTCCAAGATTAA 451
 DB 122 rgGlnTyRLeuLeuSerLysGluAlaArgGluGlyIleGlyProHisValGlnArgLeuI 142
 QY 452 TCCAACGGGATCCCTAGTCTCTGTCCTGCAATCTCCCTGGAATACTCCCTGTACCGGTTA 511
 DB 142 leGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValA 162
 QY 512 GAAAGCCTGGGACTAATGACTATCGACCAAGTACAGACTTGAGAGAGTCAATAAAGCGG 571
 DB 162 rgLysProGlyThrAsnAspTyRArgProValGlnAspLeuArgGluValAsnLysArgV 182
 QY 572 TGCAGGATATACACCCACAGTCCCGAACCTTATAACTCTTGTGCTCTCCACCCC 631
 DB 182 alGlnAspIleHisProThrValProAsnProTyRAsnLeuLeuCysAlaLeuProProG 202
 QY 632 AACGGAGCTGGTATATACGATTGGAATTTAAAGGATGCCTTCTTCGCTGAGATTACACC 691

DB 202 lnArgSerTrpTyRThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisP 222
 QY 692 CCCTAGCCCAACCACTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAGAACCGGGC 751
 DB 222 roThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyAlaGlyArgThrGlyG 242
 QY 752 AGCTCACCTGAGCCCGACTGCCCAAGGGTTCAGAAACTCCCGACCATCTTTTACCAAG 811
 DB 242 lnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluA 262
 QY 812 CCCTACACAGAGACTGCGCAACTTCAGGATCCCAACACCTCTCAGGTGACCCCTCTCCAGT 871
 DB 262 laLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnT 282
 QY 872 ACGTGGATGACTGCTCTTGGCGGAGCCACCAACAGGACTGCTTAGAGAGCCAGGAAG 931
 DB 282 yrValAspAspLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGly----- 299
 QY 932 CACTACTGCTGGAATTCCTGACCTAGCTACAGAGCTCTGCTAAGAGCCCGACAGATT 991
 DB 300 --LeuLeuLeuGluLeuPheAspLeuGlyTyRArgAlaSerAlaLysLysAlaGlnIleC 319
 QY 992 GCAGGAGAGAGGTAAACATACCTTGGGGTACAGTTTG---CGGGACGGGACGCGA---TGCG 1045
 DB 319 ysArgArgGluAlaThrGlnLeuGlyValGlnValCysGlyAlaGlyGlnSerAspTrpL 339
 QY 1046 TGACG---GAGGCGACGGAAGAAACTGTA---GTCCAGATACCGGCCCAACACACGCCA 1099
 DB 339 euThrGlyLysAlaArgLysLysThrValGlnProLysIleGlyProProThrThrAlaL 359
 QY 1100 AACAA---ATGAGAGAGTTTTTGGGGACA---GCTGGATTTTGCAGACTGGATCCGG 1153
 DB 359 ysGlnValValArgGluPhePheGlyAlaGlnValGlyPheCysArgLeuTrpIleProG 379
 QY 1154 GGTTTGGGACTTAGCAGCCCCACTCTACCCGCTAAACCAAGGAGGGAATTTCTCCT 1213
 DB 379 lyPheAlaThrLeuAlaAlaProLeuTyRProLeuThrLysGluLysGlyGluPheSerT 399
 QY 1214 GGGCTCTCTGACACACAGAGGCATTTGATGCTATCAAAAGGCCCTGCTGAGCGCACCTG 1273
 DB 399 rpAlaLeuGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeuSerSerAlaProA 419
 QY 1274 CTCTGGCCCTCCCTGACGTAAACCTTTACCTTTATGCTGATGAGGTGAAGGAG 1333
 DB 419 laLeuAlaLeuProAspValLeuLysProPheThrLeuTyRValAspGluArgLysGlyV 439
 QY 1334 TAGCCCGGGAGTTTAAACCCAAACCCCTAGGACCATGGAGAAGACCTGTCCCTACCTGT 1393
 DB 439 alAlaArgGlyValLeuThrGlnIleLeuGlyProTrpArgProValAlaTyRLeuS 459
 QY 1394 CAAAGAGCTCGATCTCTGTAGCTAGTGGTGGCCCATATGCTCAAGGCTATCGCAGCTG 1453
 DB 459 erLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLysAlaIleAlaVal 479
 QY 1454 TGGCCATACCTGCTCAAGAGCGCTCACAATTTGACTTTGGGACAGAAATATACTGTAATAG 1513
 DB 479 alAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleA 499
 QY 1514 CCCCCCATGCAATTGGAGAAACTCGTTGGCAGCCCCCAGACCCGATGGATGACCAACGCC 1573
 DB 499 laProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrpMetThrAsnAlaA 519
 QY 1574 GCATGACCCACTATCAAGCCTGCTTCTCAGAGAGGGTCAAGTTCGCTCCACGAGCG 1633
 DB 519 rgMetThrHisTyRHisGlnSerLeuLeuLeuThrGluArgValThrPheAlaProProAlaA 539
 QY 1634 CTCTCAACCCCTGGCACCCTCTCTGCTGAAGAGACTGATGAACCACTGACTCATGATTGCC 1693
 DB 539 laLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProValThrHisAspCysH 559
 QY 1694 ATCAACTATTGATTGAGGAGACTGGGTTCGCAAGGACCTTTACAGACATA---CGCTGA 1750

Db 559 is----LeuLeuIleGluThrGlyValArgLysAspLeuThrAspIleProProLeuT 578
Qy 1751 CTGAGAGAGTCTAACCTGGTTCTACTGACGGAAGCAGCTATGCTGGTGAAGCTAAGAGA 1810
Db : : : : :
Db hrGlyLysMetLeuThrTrpPheThrAspGlySerSerTyValValGluGlyLysSerM 598
Qy 1811 TGGCTGGGGCGGCGTGTGACGGGACCCGACGATCTGGCCAGCGACCTGCCGGAAG 1870
Db 598 etAlaGlyProProValValThrGlyThrArgThrIleTrpAlaSerSerLeuProGluG 618
Qy 1871 GAACTTCAGCAAAAGGCTCAGCTCATGGCCCTCAGCAAGCTTTGGCGCTGCCGAAG 1930
Db 618 lyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluG 638
Qy 1931 GGAATCCATAAACATTATACGACAGCAGGTATGCTTTGCGACTGCACACGTACATG 1990
Db 638 lyLysSerIleAsnIleTyThrAspSerArgTyAlaPheAlaThrAlaHisValHisG 658
Qy 1991 GGGCCATCTATAAACAAGGGGTGCTTACTCTCAGCAGGGGAGGAAATAAAGAACAAAG 2050
Db 658 lyAlaIleTyLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIleLysThrLysG 678
Qy 2051 AGGAATTCCTAAGCCTATTAGAAGCGGTACATTTACCANAAGGCTAGCTATTATACACT 2110
Db 678 luGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAlaIleIleHisC 698
Qy 2111 GTCCTGGACATCAGAAAGCTAAAGATCTCATATCAGAGGAAACAGATGGCTACCGGG 2170
Db 698 ysProGlyHisGlnLysAlaLysAspProIleSerArgGlyAsnGlnMetAlaAspArgV 718
Qy 2171 TTGCCAAGCGGACGCCAGGGTGTAACTTCTGCTTACCTTAATAGAAATGCCAAAGCCC 2230
Db 718 alAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProMetIleGluThrProLysAlaP 738
Qy 2231 CAGAACCCAGACGACGACACCTCAGAACCTCAGAACGTCGCAAGAGATAAAGAATAGACCACT 2290
Db 738 roGluProGlyArgGlnTyrrThrLeuGluAspTrpGlnGluIleLysLysIleAspGlnP 758
Qy 2291 TCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATCTGCCCCACA 2350
Db 758 heSerGluThrProGluGlyThrCysTyrrThrSerAspGlyLysGluIleLeuProHisL 778
Qy 2351 AAGAGGGTTAGAAATATGTCACACAGATACATCTCTAACCCCTAGGAACCTAAACACC 2410
Db : : : : :
Db ysGluGlyLeuLysTyValGlnGlnIleHisArgLeuThrHisLeuGlyThrLysHisL 798
Qy 2411 TGACGACAGTTGGTCAGAACATCCCTTATCATGTTCTGAGCTACACGAGGTGGCTCACT 2470
Db 798 euGlnGlnLeuValArgThrSerProTyrrHisValLeuArgLeuProGlyValAlaAsps 818
Qy 2471 CGGTGGTCAAACTTGTGTCCTGCCAGCTGGTTAATGCTTAATCTTCCAGAAATGCTC 2530
Db 818 erValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSerArgIleProp 838
Qy 2531 CAGGGAAGAGACTAAGGGGAAGCACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGG 2590
Db 838 roGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAspPheThrGluV 858
Qy 2591 TAAAGCCGGCTTAATACGGAACAAATACCTATTGTTGTTTGTAGACACCTTTTCAGGAT 2650
Db 858 alLysProAlaLysTyrrGlyAsnLysTyrrLeuLeuValPheValAspThrPheSerGlyT 878
Qy 2651 GGGTAGAGGCTTATCTTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAAAAATACTGG 2710
Db 878 rpValGluAlaTyrrProThrLysLysGluThrSerThrValValAlaLysLysIleLeuG 898
Qy 2711 AAGAAATTTTCCAAAGATTTGGAATACCTAAGGTAAATAGGGTTCAGACAAATGGTCCAGCTT 2770
Db 898 luGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsnGlyProAlaP 918
Qy 2771 TTGTTGCCAGGTAAAGTCAGGACTGCCCAAGATATTGGGGNTTGATTGGAACCTGCATT 2830
Db 918 heValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisC 938

Qy 2831 GTGCATACAGACCCCAAGCTCAGACAGGTTAGAGATGAATAGAACCACTTAAAGAGA 2890
Db 938 ysAlaTyrrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluT 958
Qy 2891 CCCTTACTAAATTCAGCCGAGACTGGCGTTAATGATTTGGATAGTCTCTCTGCCCTTTG 2950
Db 958 hrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpIleAlaLeuLeuProPheV 978
Qy 2951 TGCTTTTATAGGTTTGGAGAACACCCCTCGACAGATTTGGSGTGACCCCTATCAATTAATCT 3010
Db 978 alLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyrrGluLeuLeu 998
Qy 3011 ACGGGGACCCCCCTTGTGATAATTGCTTCTGTACATAGTGTGACGTGCTGCTGCTTT 3070
Db 998 yrGlyGlyProProProLeuValGluIleAlaSerValHisSerAlaAspValLeuLeuS 1018
Qy 3071 CCCAGCTTTGTTCTCTAGGCTCAAGGCATCTGAGTGGGTGAGACAACGAGCGTGGAGGC 3130
Db 1018 erGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArgAlaTrpLysG 1038
Qy 3131 AACTCCGGGAGGCTACTCAGAGGAGGAGACTTGCAGATCCCAATCGTTTCCAAGTGG 3190
Db 1038 lnLeuArgGluAlaTyrrSerGlyGluGlnValProHisArgPheGlnValG 1058
Qy 3191 GAGATTCAGTCTACTGTAGACGCCCGCTGCAGGAAACCTCGAGACTCGGTGGAAGGCC 3250
Db 1058 lyAspSerValTyrrValArgArgHisArgAlaGlyAsnLeuGluThrArgTrpLysGlyP 1078
Qy 3251 CTTATCTCGTACTTTTGACCAACACCGGCTGTGAAAGTCGAAGGAATCTCCACCTGGA 3310
Db 1078 roTyrrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIleSerThrTrpI 1098
Qy 3311 TCCATGCATCCACGTTAAACCCGCGCCACTCCCGATTCCGGGTGGAAGCCGCAAGA 3370
Db 1098 leHisAlaSerHisValLysProAlaProProAspSerGlyTrpLysAlaGluLysT 1118
Qy 3371 CTGAAAATCCCTTAAGCTTTCGCTCCATCGCGTGGTTCCCTTACTCTGTCAATAACCTCT 3430
Db 1118 hrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrrSerValAsnAsnSerS 1138
Qy 3431 CA 3432
Db 1138 er 1138

Search completed: February 14, 2006, 14:50:11
Job time : 978.589 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 12:52:14 ; Search time 91.3916 Seconds
(without alignments)
3302.856 Million cell updates/sec

Title: US-10-723-552-3_COPY_2307_5741
Perfect score: 6183
Sequence: 1 ATGGGTGCCACAGGCAACA.....CTGTCAATTAACCTTCAGAC 3435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10723552/runat_14022006_125139_12833/app.query.fasta_1
-DB=A Geneseq -QFMT=Eastan -SURFIX=rag -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN 1 1 734 @runat_14022006_125139_12833 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6038	97.7	1145	2 AAW32097	AAW32097 Miniature
2	6038	97.7	1145	4 AAB73286	AAB73286 Retroviral
3	6038	97.7	1145	8 ADS73447	ADS73447 Swine ret
4	5851.5	94.6	1194	2 AAW39272	AAW39272 Porcine r
5	4146.5	67.1	1203	2 AAW81572	AAW81572 Mus dunni
6	3818	61.7	1196	2 AAR75189	AAR75189 Osteoindu
7	3816.5	61.7	1197	4 AAB49468	AAB49468 Canine re
8	3808.5	61.6	1199	7 ADH76470	ADH76470 Murine le
9	3797.5	61.4	1199	3 AAB12994	AAB12994 MLV rever

10	3797.5	61.4	1737	3 AAB10044	AAW10044 MMLV gag-
11	3781	61.2	1224	2 AAY17947	AAY17947 MoMLV pol
12	3696	59.8	1784	2 AAR94427	AAR94427 FeLV F6A
13	3686	59.6	1784	2 AAR05898	AAR05898 Gene prod
14	3600	58.2	1079	2 AAY03163	AAY03163 MuLV reve
15	3090	50.0	1193	5 ABB79876	ABB79876 Spleen ne
16	2827	45.7	540	8 ADS73451	ADS73451 Swine ret
17	2491	40.3	672	3 AAB12457	AAB12457 Moloney m
18	2487.5	40.2	716	5 AAU74989	AAU74989 Moloney m
19	2487.5	40.2	716	8 ADN48501	ADN48501 M-MuLV re
20	2487	40.2	671	9 AEB46280	AEB46280 Murine le
21	2486	40.2	671	9 AEB46278	AEB46278 Murine le
22	2485.5	40.2	683	5 AAU74994	AAU74994 Moloney m
23	2482.5	40.2	683	5 AAU75012	AAU75012 Moloney m
24	2480.5	40.1	683	5 AAU75010	AAU75010 Moloney m
25	2480.5	40.1	683	5 AAU75017	AAU75017 Moloney m
26	2480.5	40.1	683	5 AAU74995	AAU74995 Moloney m
27	2480.5	40.1	683	5 AAU75009	AAU75009 Moloney m
28	2480.5	40.1	683	5 AAU75014	AAU75014 Moloney m
29	2480.5	40.1	683	5 AAU75011	AAU75011 Moloney m
30	2479.5	40.1	683	5 AAU74992	AAU74992 Moloney m
31	2479.5	40.1	683	5 AAU74997	AAU74997 Moloney m
32	2478.5	40.1	683	5 AAU75004	AAU75004 Moloney m
33	2477.5	40.1	683	5 AAU74990	AAU74990 Moloney m
34	2476.5	40.1	683	5 AAU75013	AAU75013 Moloney m
35	2476.5	40.1	683	5 AAU74991	AAU74991 Moloney m
36	2476.5	40.1	683	5 AAU74998	AAU74998 Moloney m
37	2476.5	40.1	683	5 AAU75008	AAU75008 Moloney m
38	2476.5	40.1	683	5 AAU75020	AAU75020 Moloney m
39	2476.5	40.1	683	5 AAU74999	AAU74999 Moloney m
40	2418.5	39.1	1183	8 ADW50751	ADW50751 Human pro
41	2221.5	35.9	1069	4 AAB73284	AAB73284 Defective
42	2051	33.2	401	2 AAW32092	AAW32092 Porcine r
43	2051	33.2	401	4 AAB73283	AAB73283 Defective
44	2051	33.2	401	8 ADS73442	ADS73442 Swine ret
45	2013	32.6	471	5 ABP47783	ABP47783 Protein #

ALIGNMENTS

RESULT 1
AAW32097
ID AAW32097 standard; protein; 1145 AA.
XX
AC AAW32097;
XX
DT 27-AUG-2003 (revised)
DT 09-FEB-1998 (first entry)
XX
DE Miniature swine retrovirus POL protein.
XX
KW Retrovirus; porcine; POL protein; xenotransplantation; infectious;
KW provirus; organ transplant; donor; activated virus; PCR.
XX
OS Pig endogenous retrovirus.
XX
FH Key Location/Qualifiers
FT Protein 1..1145
FT /label= POL_protein
XX
FN WO9721836-A1.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96WO-US019680.
XX
PR 14-DEC-1995; 95US-00572645.
XX
PA (GCHO) GEN HOSPITAL CORP.
XX
PI Fishman JA;
XX
DR WPI; 1997-332804/30.

DR N-PSDB; AAT74884.
 XX New nucleic acid from porcine retroviruses - used for detecting viruses
 PT in transplant or other tissue and for assessing risk of transmitting
 PT infection to graft recipient.
 XX
 XX Claim 22; Fig 3; 128pp; English.
 PS
 CC This is a porcine retrovirus from miniature swine containing the coding
 CC region for a putative viral POU protein. This sequence and PCR fragments
 CC generated from the sequence (see AAT74812-T74882) could be used to screen
 CC organs for porcine retroviruses prior to xenotransplantation.
 CC Transplantation can increase the likelihood of retroviral activation if
 CC intact and infectious proviruses are present. The porcine retroviral
 CC sequence can be used to generate probes to determine the level (e.g. copy
 CC number) of intact (i.e. potentially replicating) porcine provirus
 CC sequences in a strain of xenograft transplantation donors. It can be used
 CC to detect mutations, genetic lesions or viral recombinants and also to
 CC determine the histological localisation of activated retrovirus. Using
 CC Polymerase Chain Reaction DNA Quantitation (PQ) on blood mononuclear
 CC cells, infectivity titration and susceptibility testing can be performed.
 CC Ultimately animal donors without intact porcine retroviral sequences or a
 CC lower copy number of viral elements could be selected. (Updated on 27-AUG
 CC -2003 to correct OS field.)
 XX
 SQ Sequence 1145 AA;

Alignment Scores:

Pred. No.: 0 Length: 1145
 Score: 6038.00 Matches: 1145
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.7% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3_COPY_2307_5741 (1-3435) x AAW32097 (1-1145)

QY 1 ATGGGTGCACGGGACACACAGTATCATCTGGACTACCCGACAGACAGTTCAGTTGGGA 60
 DB 1 MetGlyAlaThrGlyGlnGlnGlnProTrpThrThrArgGthrValAspLeuGly 20
 QY 61 GTGGGACGGGTAAACCCATCTGCTGCTATACCTGAGTGCACAGCACCCCTCTTAGGT 120
 DB 21 ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly 40
 QY 121 AGAGACTTATTGACCAAGATGGGAGCAAAATTTCTTTTGAACAGGGAACCCAGAGTG 180
 DB 41 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal 60
 QY 181 TCTGCAATACAAACCTTACCTGCTGTGTGACCTCCCAATTAGATGACGAATATCGACTA 240
 DB 61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 80
 QY 241 TACTCTCCCTAGTAAAGCCTGTATCAAAATATACAAATTCGTGTTGGACAGTTCCTCCCA 300
 DB 81 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln 100
 QY 301 GCCTGGGAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCTCCCAAGAGTTATTCAA 360
 DB 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120
 QY 361 CTGAAGCCAGTGCACACCGATGTCTAGTCAGACAGTACCCCTTGAGTAAAGACGCTCAA 420
 DB 121 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerLysGluAlaGln 140
 QY 421 GAAGGAATTCGGCCCGCATGTCCAAAGATTAAATCCAAAGGGCATCTAGTTCTCTGCCAA 480
 DB 141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 160
 QY 481 TCTCCCTCGAATACTCCCTGCTACCGGTTAGAAAGCCTGGGACTTAATGACTATCGACCA 540
 DB 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180

QY 541 GTACAGGACTTGAGAGAGTCAATAAACCGGTGAGGATATACACCAACAGTCCCGAAC 600
 DB 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200
 QY 601 CTTTATACCTCTTTGTGTCTCCACCCCAACCGAGCTGGTATACAGTATTGGACTTA 660
 DB 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTyrThrValLeuAspLeu 220
 QY 661 AAGGATCGCTTCTTCTGCTGAGATTACACCCACTAGCAACCACTTTTTCCTCCAA 720
 DB 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
 QY 721 TGGAGAGATTCAGGTACGGGAAGAACCGGGCAGCTCACCTGCACCCGACTGCCCAAGGG 780
 DB 241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
 QY 781 TTCAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCAACTTCAGG 840
 DB 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280
 QY 841 ATCCAAACCTCTAGGTGACCTCTCCAGTACCTGAGTACCTGCTTCTGCGGGAGCC 900
 DB 281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 300
 QY 901 ACCAAACAGGACTCTTAGAAGGCACGAAGGCACCTACTCTGGAAATGTCTGACCTAGGC 960
 DB 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAspLeuGly 320
 QY 961 TACAGAGCTCTGCTAAGAGGCCAGATTTGACGAGAGAGAGTAACATCTTGGGGTAC 1020
 DB 321 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 340
 QY 1021 AGTTTGGGGACGGCGAGCGATGCTGACGAGCAGCGAAGAAACCTGTAGTCCAGATA 1080
 DB 341 SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 360
 QY 1081 CCGGCCCAACACACAGCCCAAAATGAGAGATTTTGGGGACAGCTGGATTTTGACA 1140
 DB 361 ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg 380
 QY 1141 CTGTGGATCCCGGGTTTGGACCTTACGACCCCTTACCTACCCCTAACCAAGAAAA 1200
 DB 381 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 400
 QY 1201 GGGGAATTCCTCGGCTCCTGACACACAGAGCATTTGATGCTATCAAAAGGCCCTG 1260
 DB 401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
 QY 1261 CTGAGCGCAGCTGCTGCGCCCTCCCTGACGTAAACCTTACCTTTATGTGGAT 1320
 DB 421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440
 QY 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCAACCTTAGGACCATGGAGAGACCT 1380
 DB 441 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 460
 QY 1381 GTCCCTACCTGTCAAAGAAGCTCGATCTGTAGCAGTGGTGGCCCATATGCTGAG 1440
 DB 461 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProIleCysLeuLys 480
 QY 1441 GCTATCGCAGCTGTGGCCATCTGTCAGGACCTGACAAATGACATTTGGGACAGAA 1500
 DB 481 AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500
 QY 1501 ATAACTGTAATAGCCCCCATGTCATTGGAGAACATCGTTCCGCGACCCCGACGATGG 1560
 DB 501 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 520
 QY 1561 ATGACCAAGCCCGCATGACCCACTATCAAGCCTGCTTCTCACAGAGAGGGTCAGCTC 1620
 DB 521 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540
 QY 1621 GCTCCACCGCCGCTCTCAACCCCTGCCACTCTTCTGCTGAAAGAGACTGATGAACCACTG 1680

Db 541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 1681 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 1740
Db 561 ThrHisaspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 580
QY 1741 ATACCGCTGACTGGAGAAGTGCTAACCTGGTCTACTGACGGAAGCAGCTATGTGGTGAA 1800
Db 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTy-ValValGlu 600
QY 1801 GGTAAAGAGATGGCTGGGGCGCGTGTGACGGGACCCACAGATCTGGCGCAGCAGC 1860
Db 601 GlyLysArgMetAlaGlyAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
QY 1861 CTGCGGGAAGGAATCTCAGCAAAAGGCTCAGGCTCATGGCCCTCAGCAAGCTTTGGG 1920
Db 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 1921 CTGCGCGGAAGGAATCATAAACATTTATACGACAGCAGGTATGCTTTGCGACTGCA 1980
Db 641 LeuAlaGluGlyLysSerIleAsnIleTyThrAspSerArgTyAlaPheAlaThrAla 660
QY 1981 CACCTACATGGGGCATCTATAAACAAGGGGTGCTTACTCTCAGCAGGAGGAAATA 2040
Db 661 HisValHisGlyAlaIleTyLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 2041 AAGAAACAAAGAGAAATCTTAAGCCTATTAGAAGCGGTACATTTTACCACAAAGGCTAGCT 2100
Db 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
QY 2101 ATTATACATGTCTCGGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATG 2160
Db 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
QY 2161 GCTGACCGGGTTCGCAAGCAGGAGCCAGGGGTGTTAACTTCTGCTCTATAATGAAATG 2220
Db 721 AlaAspArgValAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProIleIleGluMet 740
QY 2221 CCCAAAGCCCAAGAACCCAGACGACGTACACCTCTAGAAGACTGGCAAGAGATAAAAG 2280
Db 741 ProLysAlaProGluProArgArgGlnTyThrLeuGluAspTrpGlnGluIleLysLys 760
QY 2281 ATAGACAGATTCTGTGAGACTCCGGAAGGACCTGTATACCTCAGATGGGAAGGAATC 2340
Db 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyThrSerAspGlyLysGluIle 780
QY 2341 CTGCCCCACAAGNAGGGTTAGATATGTCCAACAGATACATGCTCTAAACCCACCTAGGA 2400
Db 781 LeuProHisLysGluGlyLeuGluTy-ValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 2401 ACTAAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCTACCAGGA 2460
Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyThrHisValLeuArgLeuProGly 820
QY 2461 GTGGCTGACTCGGTGGTCAAAACATTGTGTGCTCCCTGCGCAGCTGGTTAATGTAATCTCTCC 2520
Db 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY 2521 AGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCACTGGGAAGTGGAC 2580
Db 841 ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
QY 2581 TTCCTAGGTAAAGCCGGCTAATACGGGAAACAAATACCTATTGGTTTGTAGACACC 2640
Db 861 PheThrGluValLysProAlaLysTy-ValLysLysTyThrLeuLeuValPheValAspThr 880
QY 2641 TTTTTCAGGATGGGTAGAGGCTTATCTTACTTAAGAAAGAGACTTCAACCGGTGGGTAAA 2700
Db 881 PheSerGlyTrpValGluAlaTyProThrLysLysGluThrSerThrValValAlaLys 900
QY 2701 AAAATACTGGAAAGAAATTTTCCAAAGATTTCGAATACCTAAGGTAAATAGGGTCAGACAAT 2760

Db 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920
QY 2761 GGTCCAGCTTTTGTTCCTCCAGGTAAAGTCAGGAGCTGGCCCAAGATATTGGGATTCATGG 2820
Db 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY 2821 AAACCTGCATTGTGTATACAGACCCCAAGCTCAGGACACAGGTAGAGAGGATGAATAGAAC 2880
Db 941 LysLeuHisCysAlaTyArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY 2881 ATTTAAAGAGACCTTACTTAATTAATGACCGCGAGACTGGCGTTAATGATGATGACTCTC 2940
Db 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
QY 2941 CTGCGCTTTGTGCTTTTGTAGGTTAGGAACACCCCTGACACAGTTTGGGCTCAGCCCTAT 3000
Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTy- 1000
QY 3001 GAATTACTCTACGGGGGACCCCTCATTTGGTAGAAATTGCTTCTGTACATAGTGTCTGAC 3060
Db 1001 GluLeuLeuTy-ArgGlyProProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
QY 3061 GTGCTGCTTCCACAGCTTTGCTCTAGGCTCAAGGCACCTTGAGTGGGTGAGACACAGA 3120
Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY 3121 GCGTGGAGGCAACTCCGGGAGGCTACTCAGGAGGAGGAGACTTGCAGATCCACATCGT 3180
Db 1041 AlaTrpArgGlnLeuArgGluAlaTySerGlyGlyAspLeuGlnIleProHisArg 1060
QY 3181 TTCACAGTGGAGATTCAAGTCTAGACCGCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 3240
Db 1061 PheGlnValGlyAspSerValTy-ValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
QY 3241 TGAAGGGCCCTTATCTCGTACTTTTGACACACACCGCTGGAAGTGAAGTGAAGTGAAG 3300
Db 1081 TrpLysGlyProTy-ValLeuLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
QY 3301 TCACACTGGATCCATGCATCCCGTAAACCGCGCCACCTCCCGATTCGGGTGGGAA 3360
Db 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1120
QY 3361 GCCGAAAGACTGAAATCCCTTAAAGCTTTCGCTTCCATCGCTGGTTCCTTACTCTGTC 3420
Db 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTy-SerVal 1140
QY 3421 AATAACCTCTCAGAC 3435
Db 1141 AsnAsnLeuSerAsp 1145
RESULT 2
AAB73286
ID AAB73286 standard; protein; 1145 AA.
XX AAB73286;
XX AC
XX 23-MAY-2001 (first entry)
XX DT
XX DE Retroviral protein #2 found in miniature swine.
XX KW Retrovirus; graft transplantation; xenotransplantation; miniature swine.
XX OS Unidentified.
XX US6190861-B1.
XX PD 20-FEB-2001.
XX PF 13-DEC-1996; 96US-00766528.
XX PR 14-DEC-1995; 95US-00572645.
XX (GEHO) GEN HOSPITAL CORP.

XX Fishman JA; 181 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 200
XX PI
XX DR WPI; 2001-256211/26. 601 CCTTATAAAGCTCTTGTGTCTCTCCACCCCAAGGAGCTGGTATACAGATTGGACTTA 660
XX DR N-PSDB; AAF77727. 201 ProTyrAsnLeuLeuCysAlaLeuProProGlnArgSerTrpTrpThrValLeuAspLeu 220
XX PT Assessing risk of endogenous retroviruses in clinical practice and in 661 AAGGATGCCCTCTTCTTCCTCGAGATTACACCCACTAGCCAAACCACTTTTTCCTCGAA 720
XX PT xenotransplantation, comprises using probe sequences derived from swine 221 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 240
XX PT or miniature swine retroviral genome. 721 TGGAGAGATCCAGTACGGGAAGAACCGGCGAGCTACCTGGACCCCGACTGCCCAAGGG 780
XX PS Disclosure; Fig 3; 127pp; English. 241 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 260
XX CC The present invention relates to a method for screening a cell or tissue 781 TTCAGAAGCTCCCGACCATCTTTGACGAAGCCCTCACAGAGACCTGGCCCACTTCAGG 840
XX CC for the presence or expression of a retrovirus (RV), comprising 261 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 280
XX CC contacting a target nucleic acid from the cell or tissue with a second 841 ATCCAAACACCTCAGGTGACCCCTCCTCAGTAGCTGATGACCTGTCTGGCGGGAGCC 900
XX CC nucleic acid from the present invention (e.g. AAF77727 or a fragment 281 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuAlaGlyAla 300
XX CC thereof). The method is useful for RV detection and to assess graft 901 ACCAAACAGACTGCTTAGAAGGCACGAAGGCCTACTGCTGGAATTGTCTGACCTAGGC 960
XX CC donors with active replication of known viruses. Inactive proviruses can 301 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuLeuSerAspLeuGly 320
XX CC be detected and inactivated, allowing identification and elimination of 1021 AGTTTGGCGGACGGGAGCGATGGCTGACGAGGACGAGGAGGAGGAGGAGGAGGAGGAGG 1080
XX CC potential human pathogens derived from swine in a manner not possible in 341 SerLeuArgAspGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 360
XX CC the outbred human organ donor population and is important to the 1081 CCGGCCCAACCAACAGCCAAACAAATGAGAGAGTTTGGGACAGCTGGATTGGGAG 1140
XX CC development of human xenotransplantation 361 ProAlaProThrThrAlaLysGlnMetArgGluPheLeuGlyThrAlaGlyPheCysArg 380
XX SQ Sequence 1145 AA; 1141 CTGTGGATCCGGGGTTGGACCTTAGCAGCCCACTACCCGCTAAACCAAGAAAAA 1200
381 LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrLysGluLys 400
1201 GGGGAATTCCTCGGGCTCCTGAGCACGAGGACATTTGTATGCTATCAAAAGCCCTG 1260
401 GlyGluPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 420
1261 CTGAGCGCACCTGCTCGGCCCTCCCTGACGTAACTAAACCTTTACCTTTATGTGGAT 1320
421 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 440
1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTTAGCAGCCATGGAGAGACCT 1380
441 GluArgLysGlyValAlaAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgPro 460
1381 GTCCCTACCTGTCAAAAGAGCTCGATCCTGTAGCAGTGGTGGCCCATATGCTCGAAG 1440
461 ValAlaTyrLeuSerLysLysLysLeuAspProValAlaSerGlyTyrProIleCysLeuLys 480
1441 GCTATCGCAGCTGTGGCCCATATCTGGTCAAGGACCTGACAAATTCATCTTGGGACAGAT 1500
481 AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 500
1501 ATAACTGTAATAGCCCCCATGCAATGGTAGAGACATCGTTCGGCAGCCCGGAGCGGATGG 1560
501 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 520
1561 ATGACCAAGCGCCCATGACCCACTATCAAGAGCTGCTTCTTCACAGAGAGGCTCACGTTTC 1620
521 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 540
1621 GCTCCACGAGCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACGTGATGAACAGTGG 1680

Alignment Scores:

Pred. No.: 0 Length: 1145
Score: 6038.00 Matches: 1145
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.7% Indels: 0
DB: 4 Gaps: 0

US-10-723-552-3_COPY_2307_5741 (1-3435) x AAB73286 (1-1145)

QY 1 ATGGGTGCCACAGCGGCAACACAGTATCCATGGACTACCCGAAGAACAGTTGACTGGGA 60
DB 1 MetGlyAlaThrGlyGlnGlnIleNyrProTrpThrThrArgArgThrValAspLeuGly 20
QY 61 GTGGGACGGGTAAACCACTCGTTTCTGTGTATACCTGAGTGGCCGAGCACCCCTCTTAGGT 120
DB 21 ValGlyArgValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly 40
QY 121 AGAGACTATTGACAGATCGGAGCACAAATTTCTTTGAACAGGGAACCAAGTGT 180
DB 41 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyLysProGluVal 60
QY 181 TCTGCAATAACAAACCTATCACTGTGTGACCTCCCAATTAGATGACGAATATCGACTA 240
DB 61 SerAlaAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 80
QY 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAATTCCTGGTGGAAACAGTTTCCCAA 300
DB 81 TyrSerProLeuValLysProAspGlnAsnIleGlnPheTrpLeuGluGlnPheProGln 100
QY 301 GCCTGGGAGAAAACCGGAGGATGGGTTGGCAAGCAAGTTCCTCCCAAGATTATTCAA 360
DB 101 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 120
QY 361 CTGAGGCCAGTGCACACACAGTGTCACTGACAGACTACCCCTTGAGTAAAGAGCTCAA 420
DB 121 LeuLysAlaSerAlaThrProValSerValArgGlnIleProLeuSerLysGluAlaGln 140
QY 421 GAAGGAATTCGGCCGATGTCCAAAGATTATCCAAACAGGGCATCTTAGTTCTCTGTCCAA 480
DB 141 GluGlyIleArgProHisValGlnArgLeuIleGlnGlnIleLeuValProValGln 160
QY 481 TCTCTCCGGAATATCTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATGACCA 540
DB 161 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 180
QY 541 GTACAGGACTTGAGAGAGTCAATAAACGGGTGCAGGATATACACCCCAACAGTCCCGAAC 600

Db 541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560
QY 1681 ACTCATGATTCCTCACTAATTGATTGAGGAGACTGGGGTCCGCAAGGACTTACAGAC 1740
Db 561 ThrHisaspCysHisGlnLeuLeuIleGluGluThrGlyValAaGlysAspLeuThrAsp 580
QY 1741 ATACCGCTGACTGAGAAGTCTAACTGCTGTTCACTGACGGAAGACGCTATGTGGTGAA 1800
Db 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
QY 1801 GGTAAGAGGATGGTGGGGGGGGTGGTGGACGGGACCCGACAGACTGGCCGACGAGC 1860
Db 601 GlyIysargMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
QY 1861 CTGCCGGAAGGAATCTCAGCACAAAAGGCTGAGCTCATGGCCCTCAGCGCAAGCTTTGCGG 1920
Db 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
QY 1921 CTGCCGGAAGGAATCTAATAACATTTATACGACAGCGAGTATGCTTTGCGACTGCA 1980
Db 641 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
QY 1981 CACGTACATGGGGCATCTATACAAAGGGGTGCTTACCTCAGCAGGGAGGGAATA 2040
Db 661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
QY 2041 AAGAACAAAGAGGAATTTCTAAGCCTATTAGAACCCGTACATTTACCAAAAAGGCTAGCT 2100
Db 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
QY 2101 ATTATACATGCTCTGACATCAGAAAGCTTAAGATCTCATATCCAGAGGAACACGATG 2160
Db 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
QY 2161 GCTGACCGGTTCGCAAGCAGGAGCCAGCGGTGTTAACTTCTGCTTATATAGAAATG 2220
Db 721 AlaAspArgValAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProIleIleGluMet 740
QY 2221 CCCAAAGCCCGAAGACCCAGACAGCAGTACACCTCAGAACCTGGCAAGAGATAAAAAAG 2280
Db 741 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
QY 2281 ATGACACAGTTCTGTAGATCCCGGAAGGACCTGCTATACCTCAGATGGGAAGGAATC 2340
Db 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
QY 2341 CTGCCCCACAAAGAGGTTAGATATGTCCACAGATACATGCTTAACCCACCTAGGA 2400
Db 781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
QY 2401 ACTAAACACCTGCAGCAGTGGTTCAGAACATCCCTTATCATGTTCTGAGGCTACAGGA 2460
Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
QY 2461 GTGCGTACTCGTGTGTCAAACATGTTGTGCCCTGCCAGCTGGTTAATGCTAATCTCTCC 2520
Db 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
QY 2521 AGAATGCTCCAGGGAAGAGACTAAGGGGAAGCCAGCCAGCGCTCAGTGGGAAGTGGAC 2580
Db 841 ArgMetProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 860
QY 2581 TTCACCTGAGTAAGCCGGCTAAATACGGAACAAATACCTATTGTTTTTGTAGACACC 2640
Db 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
QY 2641 TTTTTCAGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTGGTAAA 2700
Db 881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
QY 2701 AAAATACTGGAAGAAATTTTCCAGATTTGGAAATACCTAAGTAATAGGGTCAGACAAT 2760
Db 901 LysIleLeuGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920

QY 2761 GGTCCAGCTTTTGTGTCCTCAGGTAAAGTCAGGACTGCGCAAGATATTGGGATTGATTGG 2820
Db 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 940
QY 2821 AAAGTCAATTTGTGCATACAGACCCCAAGCTCAGGACGAGTACAGAGGATGAATAGAAC 2880
Db 941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
QY 2881 ATTAAAGAGACCTTACTAAATTGACCGCGGAGACTGGCTTAATGATTGATAGCTCTC 2940
Db 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
QY 2941 CTGCCCTTTGTGCTTTTGTAGGTTAGGAACACCCCTCGACAGTTTGGGTGACCCCTTAT 3000
Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
QY 3001 GAATTAATCTACCGGGGACCCCCCTTGGTAGAAATTCCTCTGTACATAGTGTCTGAC 3060
Db 1001 GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
QY 3061 GTGCTGCTTTCCAGCCCTTTGTTCTTAGGCTCAAGCCTTGTAGTGGGTGAGACAAGGA 3120
Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
QY 3121 GCCTGGAGGCAACTCCGGGAGGCTACTCAGGAGGAGGACTTGCAGATCCACATCGT 3180
Db 1041 AlaTrpArgGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnIleProHisArg 1060
QY 3181 TTCAGAGTGGAGATCAGTCTAGTTAGACGCCACCGTCAGGAAACCTCGAGACTCGG 3240
Db 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
QY 3241 TGGAGGGCCCTTATCTGCTACTTTTGCACACACCGCTGTGAAAGTCGAGGAATC 3300
Db 1081 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1100
QY 3301 TCCACCTGGATCCATGCATCCACCTTAAACCCGCGCACCTCCCGATTCCGGGTGGAAA 3360
Db 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpLys 1120
QY 3361 GCCCAAGAGACTGAAATCCCTTAAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTC 3420
Db 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
QY 3421 AATAACCTCTCAGAC 3435
Db 1141 AsnAsnLeuSerAsp 1145
RESULT 3
ADS73447
ID ADS73447 standard; protein; 1145 AA.
XX
AC ADS73447;
XX
DT 16-DEC-2004 (first entry)
XX
DE Swine retroviral pol protein.
XX
KW Swine retroviral protein; immunosuppressive; gene therapy; pol protein.
XX
OS Porcine endogenous retrovirus.
XX
PN US2004185435-A1.
XX
PD 23-SEP-2004.
XX
PF 26-NOV-2003; 2003US-00723552.
XX
PR 14-DEC-1995; 95US-00572645.
PR 13-DEC-1996; 96US-00766528.
PR 14-SEP-2000; 2000US-00661858.
XX

Db 541 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 560
Qy 1681 ACTCATGATGCCATCACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 1740
Db 561 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValAlaGlyAspLeuThrAsp 580
Qy 1741 ATACGGCTGACTGGAGAAGTCTAACTGGTTCACTGACGGAAGCAGCTATGTTGGTGA 1800
Db 581 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 600
Qy 1801 GGTAAAGAGGATGGCTGGGGCGCGTGTGTGACCGGACCCGACAGATCTGGGCCAGCAGC 1860
Db 601 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 620
Qy 1861 CTGCGCGAAGAACTTCAGCAACAAAGGCTGAGCTCATGGCCCTCAGCAAGCTTTGCGG 1920
Db 621 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 640
Qy 1921 CTGCGCGAAGAAATCCATAAACATTTATACGACAGCAGGATGCTTTGGCAGCTGCA 1980
Db 641 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 660
Qy 1981 CACCTACATGGCGCATCTATAACAAAGGGGTTGCTTACTCAGCAGGAGGGAATA 2040
Db 661 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 680
Qy 2041 AAGAACAAAGAGGAAATTTTAAGCCTATTAGAAGCCGTACATTTACCAAAAAGGCTAGCT 2100
Db 681 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaValHisLeuProLysArgLeuAla 700
Qy 2101 ATTATACATGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATG 2160
Db 701 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 720
Qy 2161 GCTGACCGGTTGCAAGCGAGCCAGCGGTCTTAACCTTCTGCTTATATAAGAAATG 2220
Db 721 AlaAspArgValAlaLysGlnAlaAlaGlnGlyValAsnLeuLeuProIleIleGluMet 740
Qy 2221 CCCAAAGCCCCAGAACCCAGACGACAGTACACCCCTAGAAGACTGGCAAGAGATAAAAAAG 2280
Db 741 ProLysAlaProGluProArgArgGlnTyrThrLeuGluAspTrpGlnGluIleLysLys 760
Qy 2281 ATAGACAGATCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATC 2340
Db 761 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerAspGlyLysGluIle 780
Qy 2341 CTGCCCCACAAAGAGGTTAGATATGTCACAGATACATCGCTAAACCCACCTAGGA 2400
Db 781 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 800
Qy 2401 ACTAAACACCTGCAGCAGTGGTCAGAACATCCCCTTATCATGTTCTGAGGCTACCAAGGA 2460
Db 801 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 820
Qy 2461 GTGCTGACTCGGTGGTCAACATGTTGTGCTGCCCTGCCAGCTGGTTAATGTAATCTCTCC 2520
Db 821 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 840
Qy 2521 AGAATGCTCCAGGAGAGACTAAGGGGAAGCCACCCAGCGCTCACTGGGAAGTGAC 2580
Db 841 ArgMetProProGlyLysArgLeuArgLysSerHisProGlyAlaHisTrpGluValAsp 860
Qy 2581 TTCCTCAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTTTTGTAGACACC 2640
Db 861 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 880
Qy 2641 TTTTCAGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCCGTGGTGTAA 2700
Db 881 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 900
Qy 2701 AAAATACGGAAGAAATTTTCCAGATTTGGATACCTAAGGTAAATAGGGTCAGACAAT 2760
Db 901 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 920

Qy 2761 GGTCCAGCTTTTGTGTCCTAGGTAAAGTACAGGACTGGCCCAAGATATTGGGATGATTGG 2820
Db 921 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyLysAspTrp 940
Qy 2821 AAAGTCAATGTGTGATACAGACCCCAAGACTCAGACAGGTAGAGAGATGAATAGAAC 2880
Db 941 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 960
Qy 2881 ATTAAAGAGACCTTTACTAAATTGACCGGAGACTGGCGTTAATGATTGGATAGCTCTC 2940
Db 961 IleLysGluThrLeuThrLysLeuThrAlaGluThrGlyValAsnAspTrpIleAlaLeu 980
Qy 2941 CTGCCCCTTGTGCTTTTGTAGGTTAGGAACACCCCTCGACAGTTTGGGCTGACCCCTAT 3000
Db 981 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1000
Qy 3001 GAATTACTCTACGGGGACCCCGCCCATTTGGTAGAATAATGCTTCTGTACATAGTCTGAC 3060
Db 1001 GluLeuLeuTyrGlyGlyProProLeuValGluIleAlaSerValHisSerAlaAsp 1020
Qy 3061 GTGCTGCTTTCCAGCCCTTTGTTCTTAGGCTCAAGGACCTTGAGTGGGTGAGACAAG 3120
Db 1021 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1040
Qy 3121 GCGTGGAGGCAACTCCCGGAGGCGCTACTCAGGAGGAGGAGACTTGCAGATCCACATCGT 3180
Db 1041 AlaTrpArgGlnLeuArgGluAlaTyrSerGlyGlyAspLeuGlnIleProHisArg 1060
Qy 3181 TTCAAAGTGGAGATTGAGTCTAGTTAGACGCCACCGTCAGGAAACCTCGAGACTCGG 3240
Db 1061 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1080
Qy 3241 TGGAAAGGCGCTTATCTCGTACTTTTACACACACCGCTGGAAGTCAAGGAATC 3300
Db 1081 TrpLysGlyProTyrLeuValLeuLeuThrProThrAlaValLysValGluGlyIle 1100
Qy 3301 TCCACCTGGATCCATGTCATCCACGTTAAACCGCGCCACCTCCCGATTCGGGGTGGAAA 3360
Db 1101 SerThrTrpIleHisAlaSerHisValLysProAlaProProProAspSerGlyTrpLys 1120
Qy 3361 GCGGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGGTGGTTCCTTACTCTGTC 3420
Db 1121 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgValValProTyrSerVal 1140
Qy 3421 AATAACCTCTCAGAC 3435
Db 1141 AsnAsnLeuSerAsp 1145

RESULT 4

AAW39272
ID AAW39272 standard; protein; 1194 AA.
XX AAW39272;
XX

DT 27-AUG-2003 (revised)
DT 19-MAY-1998 (first entry)
XX

XX Porcine retrovirus POL protein.
XX Porcine retrovirus; PoEV; POL protein; ENV protein; GAG protein; vaccine;
XX diagnosis; xenotransplantation; prophylactic; therapeutic.
XX Pig endogenous retrovirus.
XX

PN W09740167-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-GB001087.
XX
PR 19-APR-1996; 96GB-00008164.
PR 10-FEB-1997; 97GB-00002668.

XX (QONE-) Q-ONE BIOTECH LTD.
PA (IMUT-) IMUTRAN LTD.
XX Galbraith DN, Haworth C, Lees GM, Smith KT;
XX WPI; 1997-535851/49.
DR N-PSDB; AAV09700.
XX Polynucleotide encoding porcine retrovirus expression product - useful to
PT develop products for use in vaccines, diagnosis and xeno-transplantation.
XX Claim 6; Fig 3; 69pp; English.
XX This sequence represents the porcine retrovirus (PoEV) polymerase (POL)
CC protein. This protein and other porcine retroviral proteins e.g. the
CC vision core (GAG) and envelope (ENV) proteins can be used to develop
CC viral vaccines, antisense nucleic acids, ribozymes and other antiviral
CC agents. They can also be used in xeno-transplantation technology and as
CC diagnostic tools. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 1194 AA;
Alignment Scores:
Pred No.: 0 Length: 1194
Score: 5851.50 Matches: 1107
Percent Similarity: 98.4% Conservative: 17
Best Local Similarity: 96.9% Mismatches: 17
Query Match: 94.6% Indels: 1
DB: 2 Gaps: 1
US-10-723-552-3_COPY_2307_5741 (1-3435) x AAW39272 (1-1194)
QY 1 ATGGGTGCCAGGCGCAACACAGTATCATGAGTACCCGAGAACAGTTGACTTGGGA 60
DB 50 MetGlyAlaThrGlyGlnArgGlnTyrProTrpThrThrArgArgThrValAspLeuGly 69
QY 61 GTGGGCGGTAAACCCACTCGTTCTGGTATACCTGAGTGCCGAGCACCCCTCTAGGT 120
DB 70 ValGlyArgValThrHisSerPheLeuValIleProGluCySProValProLeuLeuGly 89
QY 121 AGAGACTTATTGACCAAGTGGAGACAAATTTCTTTTGAACAGGGAACCAAGATG 180
DB 90 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 109
QY 181 TCTGCAATAACAAACCTATCACTGTGTGACCTCCCAATTAGATGAGCAATATCGACTA 240
DB 110 SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyrArgLeu 129
QY 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATTCGTGTGGACAGTTCCCAAA 300
DB 130 TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGlnGlnPheProGln 149
QY 301 GCCTGGGCGAGAAACCGCAGGAGTGGTTGGCAAGCAAGTTCCCCCAACAAAGTTATTCAA 360
DB 150 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 169
QY 361 CTGAAGGCCAGTGGCCACACCAAGTGTCACTCAGACAGTACCCCTTGAGTAAAGAGCTCAA 420
DB 170 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg 189
QY 421 GAAGGAATTCGGCGCATGTCCAAAGATTATCCACAGGGCATCTAGTTCCTGTCCAA 480
DB 190 GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 209
QY 481 TCTCCCTGGAATACCTCCCTGCTACCGGTAGAAAGCCTGGAGCTTAATGACTATCGACCA 540
DB 210 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 229
QY 541 GTACAGGACTTGAGAGGTCATTAACGGGTGACAGGATATACCCCAACAGTCCCGAAC 600
DB 230 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 249

QY 601 CCTTATAACCTCTTGCTCTCCACCCAGGAGCTGGTATACAGTATTGGACTTA 660
DB 250 ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeu 269
QY 661 AAGGATCCCTTCTCTGCTGAGATTACCCCACTAGCCAAACCACTTTTGGCTTCGAA 720
DB 270 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 289
QY 721 TGGAGAGATCCAGTACGAGGAGAACCGGCGAGCTCACCTGGACCCGACTGCCCCAGGG 780
DB 290 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 309
QY 781 TTCAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCAACTTCAGG 840
DB 310 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 329
QY 841 ATCCAACACCTCAGGTGAGCCCTCTCCAGTACGTGATGACCTGCTTGGCGGGAGCC 900
DB 330 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAla 349
QY 901 ACBAACAGGACTCTTAGAAGGACGAGGACACTACTGCTGGAAATGTCTGACTAGGC 960
DB 350 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuGluLeuSerAspLeuGly 369
QY 961 TACAGAGCTCTGCTAAGAGGCCAGATTTCAGAGGAGAGGTAAACATCTTCGGGTAC 1020
DB 370 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 389
QY 1021 AGTTTGGCGGACGGGACGATGGCTGACGAGGACGCGAAGAAACCTGTAGTCCAGATA 1080
DB 390 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 409
QY 1081 CCGGCCCCAACCCACAGCAACAAATGAGAGAGTTTGGGACAGCTGGATTTTGCAGA 1140
DB 410 ProAlaProThrThrAlaLysGlnValArgGlnPheLeuGlyThrAlaGlyPheCysArg 429
QY 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTTACCCGCTAAACCAAGAAAAA 1200
DB 430 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 449
QY 1201 GGGGAATTTCTCTGGGCTCTGAGCACAGAGGCATTTGATGTATCAAAAAGGCCCTG 1260
DB 450 GlyGlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 469
QY 1261 CTGAGCCACCTGCTCTGGCCCTCCCTGACCTAACTAAACCTTTTACCTTTATGTGAT 1320
DB 470 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 489
QY 1321 GAGCGTAGGGAGTAGCCCGGGAGTTTAAACCCAAACCCCTAGGACCATGAGAGACCT 1380
DB 490 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 509
QY 1381 GTCCCTACCTGTCAAGAGAGCTCGATCTGTAGCCAGTGGTGGCCCATATGCTGTAAG 1440
DB 510 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTyrProValCysLeuLys 529
QY 1441 GCTATCGAGCTGGGCCATCTCGTCAAGAGCCCTGCACAAATTGACTTTGGGACAGAT 1500
DB 530 AlaIleAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 549
QY 1501 ATAACTGTAAATAGCCCCCATGATGGAGAAACATCGTTCCGCGAGCCCGCCAGACCGATGG 1560
DB 550 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 569
QY 1561 ATGACCAAGCCCGCATGACCCACTATCAAGCTGTCTTCTCAGAGAGGGTCAACGTTTC 1620
DB 570 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 589
QY 1621 GCTCCACCGCGCTCTCAACCTGCCACTCTCTGCTGAGAGAGCTGATGAAACCACTG 1680
DB 590 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuLeuProGluGluThrAspGluProVal 609
QY 1681 ACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGAC 1740

Db 610 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 629
 QY 1741 ATACCGCTGACTGGAGAAAGTGTAACTCGTGGTTCACTGACGGAAGCAGCTATGTGTGGA 1800
 Db 630 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSeryrValValGlu 649
 QY 1801 GGTAAAGAGATGGCTGGCGCGCGGTGTGGACCGGACCCGACGATCTGGCGCCAGCAGC 1860
 Db 650 GlyLysAspMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 669
 QY 1861 CTGCGGGAAGAACTTCAGCACAAAGGCTGAGCTCAGCGCTCACCAGCTTTGGCG 1920
 Db 670 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 689
 QY 1921 CTGGCCGAAGGAAATCATATAAATTTATACGACAGCAGCTATGCTTTGCCACTGCA 1980
 Db 690 LeuAlaGluGlyLysSerIleAsnIleThrAspSerArgTrpAlaPheAlaThrAla 709
 QY 1981 CACGTACATGGGGCCATCTATAACAAAGGGGGTTCCTTACCTCAGCAGGGAGGGAATA 2040
 Db 710 HisValHisGlyAlaIleTrpLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 729
 QY 2041 AAGAACAAGAGAAATTTCTAAGCCTATTAGAAGCGGTACATTTACCAAAAAGGTAGCT 2100
 Db 730 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 749
 QY 2101 ATTATACATCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAACACGATG 2160
 Db 750 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgLysGlnMet 769
 QY 2161 GCTACCGGGTTCGCAAGCAGCGACCCAGGGGTGTAACCTTCCTCTATTAAGAAATG 2220
 Db 770 AlaAspArgValAlaLysGlnAlaAlaGlnAlaValAsnLeuLeuProIleLeuThr 789
 QY 2221 CCCAAGCCCCAGAACCCAGCAGCAGTACACCTTAGAAGCTGGCAGAGATAAAAAG 2280
 Db 790 ProLysAlaProGluProArgArgGlnTrpThrLeuGluAspTrpGlnGluIleLysLys 809
 QY 2281 ATACACAGATTCCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGGAATC 2340
 Db 810 IleAspGlnPheSerGluThrProGluGlyThrCystyrThrSeryrGlyLysGluIle 829
 QY 2341 CTGCCCCAACAAAGAGGGTTAGAATATGTCACACAGATACATCGCTAAACCCACTAGGA 2400
 Db 830 LeuProHisLysGluGlyLeuGluTrpValGlnGlnIleHisArgLeuThrHisLeuGly 849
 QY 2401 ACTAAACACCTGCAGCGAGTGGTCAGAACATCCCTTATCATGTTCTGAGGCTACCGGA 2460
 Db 850 ThrLysHisLeuGlnGlnLeuValArgThrSerProTrpHisValLeuArgLeuProGly 869
 QY 2461 GTGGCTCAGCTCGGTGGTCAACATTTGTGTCCTGCCAGCTGGTTAATGCTATCTCTCC 2520
 Db 870 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 889
 QY 2521 AGAATGCTCCAGGGAAGACTAAGGGGAAGCCACCCAGCGCTCACTGGGAAGTGCAC 2580
 Db 890 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 909
 QY 2581 TTCTCAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTTGTTTGTAGACACC 2640
 Db 910 PheThrGluValLysProAlaLysTyrrGlyAsnLysTyrrLeuLeuValPheValAspThr 929
 QY 2641 TTTTTCAGATGGGTAGAGGCTTATCTCTACTAAGAAAGAGACTTCAACCGTGGTCTTAAA 2700
 Db 930 PheSerGlyTrpValGluAlaTrpProThrLysLysGluThrSerThrValValAlaLys 949
 QY 2701 AAAATACTGGAAGAAATTTTCCAAAGATTTGGATACCTAAGGTAATAGGTCAGACAAT 2760
 Db 950 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 969
 QY 2761 GTTCCAGCTTTGTGCCCCAGGTAAAGTCAGGAGCTGCCCAAGATATTGGGATTTGATGG 2820

Db 970 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 989
 QY 2821 AAACCTGCATTCTGTGATACAGACCCCAAGCTCAGACAGGTAGAGAGGATCAATAGAAC 2880
 Db 990 LysLeuHisCysAlaTrpArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 1009
 QY 2881 ATTAAGAGAGACCTTTACTAAATTTGACCGCGAGAGCTGCGCTTAATGATTGGATAGCTCTC 2940
 Db 1010 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeu 1029
 QY 2941 CTGCCCCTTGTGCTTTTGTAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTAT 3000
 Db 1030 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTrp 1049
 QY 3001 GAATTTACTCTACGGGGGACCCCGCTTGTAGAAATTTGTTAGTCTCTGTACATAGTCTGCAC 3060
 Db 1050 GluLeuLeuTrpGlyProProProLeuAlaGluIleAlaPheAlaHisSerAlaAsp 1069
 QY 3061 GTGCTGCTTTCCAGCCCTTTTCTTAGGCTCAAGGCACCTTGAGTGGGTGAGACCAACGA 3120
 Db 1070 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1089
 QY 3121 GCGTGGAGGCACTCCGGAGGCTTACTCAGGAGGAGAGACTTGCAGATCCCATCGT 3180
 Db 1090 AlaTrpLysGlnLeuArgGluAlaTrpSer---GlyGlyAspLeuGlnValProHisArg 1108
 QY 3181 TTCCAACTGGGAGATTCACTGTTAGACGCCCGTGCAGAAACCTCGAGACTCGG 3240
 Db 1109 PheGlnValGlyAspSerValTrpValArgHisArgAlaGlyAsnLeuGluThrArg 1128
 QY 3241 TGGAAGGGCCCTTATCTCTGTTTTCACCAACACCGCTGTGAAGTCCGAAGGAATC 3300
 Db 1129 TrpLysGlyProTrpLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1148
 QY 3301 TCCACCTGGATCCATGATCCAGTTAAACCGCGCCACCTCCCGATTTCGGGTGGAAA 3360
 Db 1149 ProThrTrpIleHisAlaSerHisValLysProAlaProProAspSerGlyTrpArg 1168
 QY 3361 GCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGGTGGTTCCTTACTCTGTC 3420
 Db 1169 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTrpSerAsn 1188
 QY 3421 AATAAC 3426
 Db 1189 AsnAsn 1190
 RESULT 5
 ID AAW81572
 AC AAW81572;
 XX 01-MAR-1999 (first entry)
 DT Mus dunni endogenous virus Pol protein.
 DE MDEV; retrovirus; packaging cell line; gene transfer; gene therapy;
 KW vector; polymerase protein; Pol protein.
 OS Mus dunni endogenous virus.
 XX WO9805038-A1.
 XX 12-NOV-1998.
 XX 08-MAY-1998; 98WO-US009452.
 XX 09-MAY-1997; 97US-0046140P.
 XX 08-MAY-1998; 98US-00075272.
 XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.
 XX Miller AD, Wolgamot G, Bonham L;

XX WPI; 1999-034718/03.
 DR N-PSDB; AAV69750.
 XX
 PT New retroviral packaging cells - containing Mus dunni endogenous virus
 PT sequences to target cells, retrovirus gag and pol genes and a
 PT heterologous gene of interest.
 XX
 FS Disclosure; Page 68-71; 85pp; English.
 XX
 CC This is the amino acid sequence of the polymerase (Pol) protein of Mus
 CC dunni endogenous virus (MDEV), as deduced from the MDEV nucleotide
 CC sequence (see AAV69750). A cultured packaging cell is claimed which
 CC produces a replication-defective retroviral vector (RDRV) particle, where
 CC the packaging cell is a vertebrate cell capable of expressing and
 CC assembling retroviral proteins, comprising: (a) a first vector encoding a
 CC retroviral envelope protein having amino acid residues MDEV that direct
 CC binding of the retroviral particle to MDEV retroviral receptors on a
 CC target cell; and (b) a second vector encoding retrovirus Gag and Pol
 CC proteins, where upon expression of the vectors in the packaging cell in
 CC the presence of a vector having a sequence of a heterologous gene of
 CC interest, a replication-defective retroviral particle is produced that
 CC binds to MDEV receptors of target cells. Also claimed are: (1) a cultured
 CC packaging cell for producing a RDRV particle; (2) methods for producing a
 CC RDRV particle comprising a heterologous gene of interest; (3) cultured
 CC packaging cell line PD223; and (4) a RDRV produced by a method as in (2).
 CC The MDEV receptor is present on a variety of cells rendering MDEV
 CC pseudotype packaging cells useful in methods of mammalian and
 CC particularly human gene transfer for gene therapy. The MDEV packaging
 CC cells are a stable and reproducible source of retroviral particles.
 CC Clones may be isolated from these populations that produce high titre
 CC virus. The packaging cell lines may be selected and cloned for other
 CC desirable properties, such as stability of in vivo growth, lack of
 CC production of helper virus, lack of reinfection by viral particles
 CC packaged in the cell, stability from genetic rearrangement and
 CC recombinational events, resistance to complement lysis, and improved
 CC ability to infect cells from higher mammals
 XX
 SQ Sequence 1203 AA;

Alignment Scores:

Pred. No.:	0	Length:	1203
Score:	4146.50	Matches:	786
Percent Similarity:	79.6%	Conservative:	125
Best Local Similarity:	68.6%	Mismatches:	213
Query Match:	67.1%	Indels:	21
DB:	2	Gaps:	7

US-10-723-552-3_COPY_2307_5741 (1-3435) x AAW81572 (1-1203)

Qy	1	ATGGGTGCCACAGGGCAACACAGATATCCATGGACTACCCGAGAACACAGTTCAGTTGGGA	60
Db	50	ValGlyAlaThrGlySerLysLeuTyrProTrpThrThrLysArgAlaLeuGlnIleAsp	69
Qy	61	GTGGGACGGTAAACCCATCTGCTTCTGCTGATACCTGAGTGCCAGCACCCCTCTTAGGT	120
Db	70	LysAsnMetValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly	89
Qy	121	AGAGACTTATTGACCAAGATGGGACACAAATTTCTTTTGAAACAGGGAAACCAAGATG	180
Db	90	ArgAspLeuLeuThrLysLeuLysAlaGlnValGlnPheThrSerSerGluGlyProGlnVal	109
Qy	181	TCTGCAATAACAAACCTATCACTGTGTGGTGGTCCCAATTAGATGACGATATCGACTA	240
Db	110	SerTrpGlyLysAlaProLeuAlaCysLeuValLeuSerThrGluGluGluTyrArgLeu	129
Qy	241	TACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATTCGTGTTGGAGACAGTTCCTCCCA	300
Db	130	HisGluGluGlnProLysGlyAlaAlaProLeuAsp---TrpValThrAlaPheProAsn	148
Qy	301	GCCTGGGAGAAACCGCAGGGATGGGTTGGCAAGCAAGTTCCTCCCAAGATTATTCAA	360
Db	149	ValTrpAlaGluGlnAlaGlyMetGlyLeuAlaLysGlnValProProValValValGlu	168

Qy	361	CTGAAGGCCAGTGCACACAGTCTCAGTCAGACAGTACCCCTTGAGTAAAGAAGCTCAA	420
Db	169	LeuLysAlaAspAlaThrProIleSerValArgGlnTyrProMetSerLysGluAlaLys	188
Qy	421	GAAGAAATTCGGCGCATGTCCAAAGATTAAATCAACAGAGGCATCTAGTTCCTGTCCAA	480
Db	189	GluGlyIleArgProHisIleArgArgLeuLeuAspGlnGlyIleLeuValAlaCysGln	208
Qy	481	TCTCCCTGGAAATCTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGACTATCGACCA	540
Db	209	SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro	228
Qy	541	GTACAGCACTTGAGAGAGGTCAATAAACGGGTGAGGATATACACCAACAGTCCCGAAC	600
Db	229	ValGlnAspLeuArgGluValAsnLysArgValLeuAspIleHisProThrValProAsn	248
Qy	601	CCTTATACCTCTTGTGTGCTCTCCACCCCAAGAGCTGGTATACAGTATTGGACTTA	660
Db	249	ProTyrAsnLeuLeuSerSerLeuProGluArgThrTrpTyrThrValLeuAspLeu	268
Qy	661	AAGATGCTTCTTCTGCTGAGATTACACCCACTAGCCACCACTTTTGTCTTCGAA	720
Db	269	LysAspAlaPhePheCysLeuArgLeuHisProLysSerGlnLeuLeuPheAlaPheGlu	288
Qy	721	TGAGAGATCCAGTACGGGAAGAACCGGCGAGCTCACCTGGACCCGACTGCCCCCAAGGG	780
Db	289	TrpArgAspProGluGlyGlyGlnThrGlyGlnLeuThrTrpThrArgLeuProGlnGly	308
Qy	781	TTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTCAGG	840
Db	309	PhelysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaProPheArg	328
Qy	841	ATCCAACACCTCAGTGAGCCCTCCTCAGTACGTGATGACCTGCTTCTGGCGGAGCC	900
Db	329	AlaGlnAsnProGlnLeuThrLeuLeuGlnTyrValAspAspLeuLeuIleAlaAla	348
Qy	901	ACCAACAGGACTGCTTAGAAGCACGAAGCAGCTACTGCTGGAATTTGCTGACTAGGC	960
Db	349	SerLysGluLeuCysGlnGlnGlyThrGluArgLeuLeuThrGluLeuGlyAsnLeuGly	368
Qy	961	TACAGAGCCTCTGCTAAGAAGCCCGAGATTTTGCAGAGAGAGAGTAAACATACCTGGGTAC	1020
Db	369	TyrArgValSerAlaLysLysAlaGlnIleCysGlnThrGluValIleTyrLeuGlyTyr	388
Qy	1021	AGTTTGGGACGGGACGAGTGGCTGACGAGGACCGAAGAAACCTGATGCTCAGATA	1080
Db	389	ThrLeuArgGlyGlyLysArgTrpLeuThrGluAlaArgLysLysThrValMetMetIle	408
Qy	1081	CCGGCCCCAACACAGCCAAACAAATGAGAGAGTTTTGGGACAGCTGGGATTTTGACA	1140
Db	409	ProProThrThrProArgGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg	428
Qy	1141	CTGTGGATCCCGGGTTTGGACCTTACGAGCCCCACTCTACCCGCTAACCAAGAAAA	1200
Db	429	LeuTrpIleProGlyPheAlaThrLeuAlaProLeuTyrProLeuThrArgGluGly	448
Qy	1201	GGGGAATTCCTCGGCTCTGAGCACAGAGCATTGTGCTATCAAAAAGCCCTG	1260
Db	449	IleProPheGluTrpLysGluHisGlnArgAlaPheGluAlaIleLysSerSerLeu	468
Qy	1261	CTGAGCGCCTCTCTGGCCCTCCCTGACGTAATAACCCCTTTTACCTTTATGTGGAT	1320
Db	469	MetThrAlaProAlaLeuAlaLeuProAspLeuThrLysSerPheValLeuTyrValAsp	488
Qy	1321	GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCAACCCCTAGGACCATGAGAGACCT	1380
Db	489	GluArgAlaGlyIleAlaArgGlyValLeuThrGlnAlaLeuGlyProTrpLysArgPro	508
Qy	1381	GTCCCTTACCTGTCAAAGAGCTCGATCCTGTACCGAGTGGTGGCCCATATGCTGAG	1440
Db	509	ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProThrCysLeuLys	528

XX RFB retrovirus; gag; pol; env; osteogenesis; osteoinductive protein;
KW bone development; osteoporosis; gene therapy; polymerase;
KW reverse transcriptase.
XX Murine leukemia virus.
XX DE4411718-A1.
XX 12-OCT-1995.
XX 05-APR-1994; 94DE-04411718.
XX 05-APR-1994; 94DE-04411718.
XX (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEIT.
XX Schmidt J, Gimbel W, Strauss P, Erfle V, Pedersen FS, Pedersen L;
PI Oestergaard M;
XX WPI; 1995-352078/46.
XX N-ESDB; AAQ94266.
XX RFB-14 retrovirus genome - and prodn. of osteo-inductive proteins.
XX Claim 14; Fig 1; 46pp; German.
XX The full-length proviral genomic sequence of retrovirus RFB-14 has been
CC determined. The virus codes for an osteoinductive protein, although the
CC precise location of the coding region has not yet been identified. The
CC virus may be useful in gene therapy of bone growth disorders such as
CC osteoporosis. The present sequence is that of the viral pol gene product.
CC (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 1196 AA;

Alignment Scores:
Pred. No.: 0 Length: 1196
Score: 3818.00 Matches: 721
Percent Similarity: 76.9% Conservative: 162
Best Local Similarity: 62.8% Mismatches: 251
Query Match: 61.7% Indels: 14
DB: 2 Gaps: 9

US-10-723-552-3_COPY_2307_5741 (1-3435) x AAR75189 (1-1196)

QY 4 GGTGCCAGGGCAACACAGTATCCATGGACTACCCGAAGAACAGTTGACTTGGAGTG 63
DB 51 GlyAlaThrGlyGlyLysArgTyrArgTrpThrThrAspArgLysValHisLeuAlaThr 70
QY 64 GGACGGGTAAACCCTCGTTTCTGGTCACTAGTGGTCCAGCACCCCTCTTAGGTAGA 123
DB 71 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 90
QY 124 GACTTATTGACAGATGGGAGCCAAATTTCTTTTGA---CAAGGAACACAGAGTG 180
DB 91 AspleuLeuThrLysLeuLysValGlnIleHisPheGluGlySerGlyAlaGlnValVal 110
QY 181 TCTGCAATAACAACACCTATCATCTGTGTGGTCCCAATTAGATGACGAATATCGACTA 240
DB 111 GlyProLysGlyGlnProLeuGlnValLeuThrLeuAsnLeuGluAspGluTyrArgLeu 130
QY 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATA---CAATTCTGGTTGGAAACAGTTTCCC 297
DB 131 TyrGluThrSerAlaGluProGluAlaSerProGlySerThrTrpLeuSerAspPhePro 150
QY 298 CAACCTGGGAGAAACCCGAGGTGGTTGGCAAGCAAGTTCCCCCAAGTATT 357
DB 151 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 170
QY 358 CAATGAAGGCGGAGTGCACACCACTGTCAGTCAGTACAGACAGTACCCCTTGAGTAAAGAGCT 417
DB 171 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrProMetSerGlnGluAla 190

QY 418 CAAGAAGGAATTCGGCGCATGTCCRAAGATTATCCACAGGGCATCTCTAGTTCCTGTC 477
DB 191 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 210
QY 478 CAATCTCCCTGGAATACTCCCTGTACTACCGGTAGAAAAGCTGGGACTAATGACTATCGA 537
DB 211 GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 230
QY 538 CCAGTACAGGACTTGAGAGAGGTCAATAAACGGGTGCAGGATATACACCCACAGTCCCG 597
DB 231 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 250
QY 598 AACCTTTATAACCTCTGTGTCTCTCCACCCCAACGGAGCTGTATACAGTATTGGAC 657
DB 251 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisArgTrpTyrThrValLeuAsp 270
QY 658 TTAAGAGATCGCTTTCTCTGCTGAGATTACACCCCTAGCCCAACCACTTTTTCCTTC 717
DB 271 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 290
QY 718 GAATGGAGATCCAGGTACGGGAAGAACCGGGAGCTCACCTGGAGCCCGACTGCCCAA 777
DB 291 GluTrpArgAspProGlyMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 310
QY 778 GGGTTCAAGAACTCCCGGACCATCTTTGACGAGCCCTACACAGACCTGGCCCACTTC 837
DB 311 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaGlyPhe 330
QY 838 AGGATCCAAACACCTCAGGTGACCTCTCCTCAGTACGTGGATGACCTGCTTCTCGCGGGA 897
DB 331 ArgIleGlnHisProAspLeuLeuLeuGlnTyrValAspAspLeuLeuLeuAlaAla 350
QY 898 GCCACCAACAGGACTGCTTAGAAGGACGAGGACACTACTGCTGGAATTTCTGACCTA 957
DB 351 ThrSerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAspLeu 370
QY 958 GGCTACAGAGCTCTGCTTAGAAGGCCCGACATTTGACGAGAGAGAGATACTTGGGG 1017
DB 371 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 390
QY 1018 TACAGTTTGGGGACGGGCGACGATGCTGACGAGGAGCAGGAAAGAACTGTAGTCCAG 1077
DB 391 TyrLeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly 410
QY 1078 ATACGGGCCCCAACACAGCCAAACAAATGAGAGAGATTTTGGGAGAGCTGGATTTCG 1137
DB 411 GlnProIleProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 430
QY 1138 AGACTGTGGATCCCGGGTTTGGACCTTAGCAGCCCTTACCGCTCTACCGCTAACCAAGAA 1197
DB 431 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 450
QY 1198 AAAGGGGAATTTCTCTGGGCTCTGAGCACCAGAAAGCATTTGATGTCTATCAAAAGGCC 1257
DB 451 GlyThrLeuPheAsnTrpGlyProAspGlnGlnLysAlaTyrGlnGluIleLysGlnAla 470
QY 1258 CTGCTGAGCGCACCTGCTCTGGCCCTCCTGACATACTAAACCTTTTACCTTTATGTG 1317
DB 471 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 490
QY 1318 GATGAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCCAACCCCTAGGACCATGAGGAGA 1377
DB 491 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArgArg 510
QY 1378 CCTGCTGCTACCTGTCAAAGAAGCTCATCTCTAGCTAGCTGGTGGCCCATATGCTG 1437
DB 511 SerValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTyrProCysLeu 530
QY 1438 AAGCTATCGAGCTGTGGGCCATATCTGTCAGGACCGCTGACAAATTGACTTTGGGACAG 1497
DB 531 ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln 550


```

XX WO200070024-A2.
XX
XX 23-NOV-2000.
XX
XX 17-MAY-2000; 2000WO-EP004467.
XX
XX 17-MAY-1999; 99EP-00401192.
XX
XX 18-MAY-1999; 99EP-00401199.
XX
XX (FRSA-) ETAB FR DU SANG.
XX
XX Rigal D, Ghernati I, Corbine A, Darlix J;
XX
XX WPI; 2001-016224/02.
XX
XX New infectious retrovirus isolated from a canine cell line, useful for
XX producing medicaments to treat autoimmune diseases, hematopoietic
XX malignancies or malignant tumors and in diagnosis and gene therapy.
XX
XX Disclosure; Fig 17; 131pp; English.
XX
XX The present invention relates to a retrovirus of type C morphology, which
XX sediments in a sucrose gradient at a density of 1.16-1.18 g/1. The
XX retrovirus is infectious for canine cells and belongs to the oncovirinae
XX group. The present sequence is the pol protein of the retrovirus of the
XX present invention, Eco (see A489138). The retrovirus can be included in
XX pharmaceutical compositions or medicaments to treat autoimmune diseases,
XX haematopoietic malignancies or malignant tumours, especially in humans.
XX The retrovirus can also be used in gene therapy to introduce a transgene
XX into an animal, especially a human
XX
XX Sequence 1197 AA;
SQ

```

```

Alignment Scores:
Pred. No.: 0 Length: 1197
Score: 3816.50 Matches: 721
Percent Similarity: 77.2% Conservative: 166
Best Local Similarity: 62.8% Mismatches: 247
Query Match: 61.7% Indels: 15
DB: 4 Gaps: 10

```

US-10-723-552-3_COPY_2307_5741 (1-3435) x AAB49468 (1-1197)

```

191 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 210
    :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
478 CAATCTCCTGGATATCTCCCTGCTACCGGTAGAAAGCTGGAGCTAATGACTATCGA 537
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
211 GlnSerProTrpAsnThrProLeuLeuProValLysProGlyThrAsnAspTrpArg 230
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
538 CCAGTACAGGACTTGAGAGAGGTCAATAAAGCGGTGAGGATATACACCAAGTCCCG 597
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
231 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 250
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
598 AACCTTTATAACCTCTTTGTGTCTCCCAACCGAGCTGTGTATACAGTATTGGAC 657
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
251 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisArgTrpTyrThrValLeuAsp 270
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
658 TTAAGGATGCTCTTCTGCTGAGATTACACCCACTAGCAACCACTTTTGCCTTC 717
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
271 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 290
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
718 GAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGACCCGACTGCCCAA 777
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
291 GluTrpArgAspProGlyMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 310
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
778 GGGTTCAAGAACTCCCGCACTCTTTGACGAAGCCCTACACAGAGACTGGCCAACTTC 837
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
311 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 330
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
838 AGGATCAACACCTCCTAGGTGACCTCTCTCAGTACGTGGATGACCTGCTCTGGCGGA 897
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
331 ArgIleGlnHisProAspLeuIleLeuLeuGlnTyrValAspAspIleLeuLeuAla 350
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
898 GCACCAACAGGAGCTCTTAGAAGCAGCAAGCACTACTGCTGGAATCTCTGACCTA 957
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
351 ThrSerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuLeuThrLeuGlyAsnLeu 370
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
958 GGCTACAGAGCTCTGCTAAGAGCCCGCCAGATTGTCAGAGAGAGGTAAACATACTGGGG 1017
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
371 GlyTyrArgAlaSerAlaLysLysAlaGlnLeuCysGlnLysGlnValLysTyrLeuGly 390
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1018 TACAGTTTGGGACGGCGAGCGATGGCTGACGGAGCGACGGAAGAAACTGTAGTCCAG 1077
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
391 TyrLeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly 410
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1078 ATACGGCCCCCAACACAGCCCAACAAATGAGAGAGATTTTGGGAGCAGCTGGATTTCG 1137
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
411 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 430
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1138 AGACTGTGGATCCCGGGGTTTGGCACTTAGCAGCCCACTCTACCCGCTAACCAAGAA 1197
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
431 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 450
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1198 AAAGGGGAATCTCTGGGCTCTGAGCACCAGAGCAATTTGATGTCTATCAAAAGGCC 1257
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
451 GlyThrLeuPheAsnTrpGlyProAspGlnLysAlaTyrGlnGluIleLysGlnAla 470
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1258 CTGCTGAGCGCACCTGCTCTGGCCCTCCCTCAGCTAACTAAACCTTTTACCCTTTATGTG 1317
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
471 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 490
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1318 GATGAGCGTAAAGGAGTAGCCCGGGAGTTTAAACCAAAACCTTAGGACCATGAGAAAGA 1377
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
491 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArgArg 510
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1378 CCTGTGCGCTACCTGTCAAGAACTCGATCTCTAGCCAGTGGTGGCCCATATGCTGTG 1437
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
511 ProValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTrpProCysLeu 530
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1438 AAGGCTATCGCAGCTGTGGCCATACTGTCAAGGACGCTGACAAATTCATCTTTGGGACAG 1497
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
531 ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln 550
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1498 AATAAATCTGTAATAGCCCCCATGTCATTTGAGAACATCGTTTCGGCAGCCGCCAGACGA 1557
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```


Db 551 ProLeuValIleLeuAlaProHisAlaValGluAlaLeuValLysGlnProProAspArg 570
QY 1558 TGGATGACAAAGCCCGCATGACCCACTATCAAGCCCTGCTTC- --ACAGAGAGGCTC 1614
Db 571 TrpLeuSerAsnAlaArgMetThrHisTyrGlnAlaMetLeuLeuAspThrAspArgVal 590
QY 1615 ACGTTCGCTCCACAGCCGCTCTCAACCTGCGCACTCTTCCTGCGTGAAGAGACTGTAA 1674
Db 591 GlnPheGlyProValValAlaLeuAsnProAlaThrLeuLeuPro- --LeuProGluGlu 609
QY 1675 CCAGTGACTCATGATTGGCCATCACTATTGATTGAGGAGACTGGGGTCCGCAAGACCTT 1734
Db 610 GlyAlaProHisAspCysLeuGluIleLeuAlaGluThrHisGlyThrArgProAspLeu 629
QY 1735 ACAGACATACCGCTGACTGGAGAAAGTGTAACTGGTTCACTGACGGAAGCAGCTATGTG 1794
Db 630 ThrAspGlnProIleProAspAlaAspHisThrTrpTyrThrAspGlySerSerPheLeu 649
QY 1795 GTGGNAGGTAAAGAGATGGCTGGGGCGCGGTGTGGACGGGACCGCACGATCTGGGCC 1854
Db 650 GlnGluGlyGlnArgLysAlaGlyAlaAlaValThrThrGluThrGluValIleTrpAla 669
QY 1855 AGCAGCCTGCCGGAAGGAATTCAGACAAAGAGCTGAGTCATGCGCCCTCAGCAAGCT 1914
Db 670 ArgAlaLeuProAlaGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAla 689
QY 1915 TTGGCGCTGGCGGAAGGAATCCATAAACTATATACGACAGCAGGTATGCTTTGCG 1974
Db 690 LeuLysMetAlaGluGlyLysArgLeuAsnValTyrThrAspSerGlnTyrAlaPheAla 709
QY 1975 ACTGCACATGATGGGGCCCATCTATAAACAAGGGGGTTGCTTACTCAGCAGGAGG 2034
Db 710 ThrAlaHisIleHisGlyGluIleTyrArgArgArgGlyLeuLeuThrSerGluGlyArg 729
QY 2035 GAAATAAGAACAAAGAGGAATCTAAGCCCTATTAGACCGGTACATTTACCAAAAAG 2094
Db 730 GluIleLysAsnLysSerGluIleLeuAlaLeuLysAlaLeuPheLeuProLysArg 749
QY 2095 CTAGCTATTATACACTGCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAC 2154
Db 750 LeuSerIleIleHisCysLeuGlyHisGlnLysGlyAspGlyAlaGluAlaArgGlyAsn 769
QY 2155 CAGATGCTGACCGGGTTGCCAAGCAGGAGCC-----CAGGGTGTAACTT 2202
Db 770 ArgLeuAlaAspGlnAlaAlaArgGluAlaAlaIleLysThrProProAspThrSerThr 789
QY 2203 CTGCTTATATAGAAATGCCAAAGCCCCAGAACCCAGACAGTACACCTAGAGAC 2262
Db 790 LeuLeuIleGluAspSerThrProTyrThrProAlaTyrPheHisTyrThrGluThrAsp 809
QY 2263 TGGCAAGAGATAAAAAAGATAGACCACTTCTCTGAGACTCCGGGAAGGACCTGCTATACC 2322
Db 810 LeuLysLysLeuArgGluLeuGlyAlaThrTyrAsnGlnSerLysGly---TyrTrpVal 828
QY 2323 TCAGATGGGAAGGAATCTGCCCCCAACAAGAGGGTTAGAAATATGTCACACAGATACAT 2382
Db 829 PheGlnGlyLysProValMetProAspGlnPheValPheGluLeuLeuAspSerLeuHis 848
QY 2383 CGTCTAACCCACTAGAACTAAACACTCGACGAGTGTGTC-----AGAACATCCCT 2436
Db 849 ArgLeuThrHisIleLeuGlyTyrGlnLysMetLysAlaLeuLeuAspArgGlyGluSerPro 868
QY 2437 TATCATGTTCTGAGGCTACCAAGGAGTGGCTGACTCGGTGCTCAAAACATTGTGCTGCTGC 2496
Db 869 TyrTyrMetLeuAsnArgAspLysThrLeuGlnTyrValAlaAspSerCysThrValCys 888
QY 2497 CAGCTGGTGAATGCTTAATCTTCCAGAAATGCCCTCCAGGGAAGAGACTAAAGGGGAAGCCAC 2556
Db 889 AlaGlnValAsnAlaSerLysAlaLysIleGlyAlaGlyValArgValArgGlyHisArg 908
QY 2557 CAGGGCGCTCACTGGGAGTGGACTTCACGTAGGTAAAGCCGGCTAAATACGGAACAAA 2616
Db 909 ProGlySerHisTrpGluIleAspPheThrGluValLysProGlyLeuTyrGlyTyrLys 928

QY 2617 TACCTATTGGTTTGTAGACACCTTTTCAGATGGGTAGAGGCTTATCTTACTTAAGAAA 2676
Db 929 TyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaPheProThrLysArg 948
QY 2677 GAGACTTCAACCGTGGTGGCTAAAAAAATACTGGAAGAAAATTTTTCGAAGATTTGGAATA 2736
Db 949 GluThrAlaArgValValSerLysLysLeuLeuGluIlePheProArgPheGlyMet 968
QY 2737 CCTAAGGTAAATAGGTCAGCAATGGTCCAGCTTTTGTGTCGCCAGGTAAAGTCAGGACTG 2796
Db 969 ProGlnValLeuGlySerAspAsnGlyProAlaPheThrSerGlnValSerGlnSerVal 988
QY 2797 GCCAAGATATTGGGGAATGATTGGAAACTGCATTGTGCATACACAGCCCAAGCTCAGGA 2856
Db 989 AlaAspLeuLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSerGly 1008
QY 2857 CAGGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACTAAATGAGC- --GCGGAG 2913
Db 1009 GlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuAlaAla 1028
QY 2914 ACTGGCGTTAATGATTGATAGTCTCTGCTGCCCTTTGTGCTTTTAGGTTAGGAACACC 2973
Db 1029 ProGlyThrArgAspTrpValLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThr 1048
QY 2974 CCTGGACAGTTTGGGCTGACCCCTATGAATTAATCTCTACGGGGGACCCGCCCATTTGTA 3033
Db 1049 ProGlyProHisGlyLeuThrProTyrGluIleLeuTyrArgAlaProProLeuVal 1068
QY 3034 GAAATTCTCTGTACATAGTGTGCTGCTGCTTTCCAGCGCTTTGTTCTCTAGGCTC 3093
Db 1069 AsnPheHisAspProAspMetSerGluLeuThrAsnSerProSerLeuGlnAlaHisLeu 1088
QY 3094 AAGCACTTCAGTGGGTGACACACAGCGTGGAGGCACTCCGGGAGGCTTACTCAGGA 3153
Db 1089 GlnAlaLeuGlnThrValGlnArgGluIleTrpAsnProLeuAlaGluAlaTyrArgAsp 1108
QY 3154 GGAGGAGACTTGCAG- --ATCCACACATCGTTTCCAGTGGGAGATTTCAGTCTAGTTAGA 3210
Db 1109 GlnLeuAspGlnProValIleProHisProPheArgIleGlyAspSerValTrpValArg 1128
QY 3211 CGCCACCGTGCAGGAACCTCGAGACTCGGTGGAAGGCGCTTATCTGCTACTTTTGACC 3270
Db 1129 ArgHisGlnThrLysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThr 1148
QY 3271 ACACCAACGCTGTGAAGTCGAAGGATCTCCACCTGGATCCATGTCATCCACGTTAAA 3330
Db 1149 ThrProThrAlaLeuLysValAspGlyIleSerAlaTrpIleHisAlaHisValLys 1168
QY 3331 CCGGCG- --CCACCTCCCGATTCCGGGTGGAAAGCCGAAAGACTGAAAAATCCCTT 3384
Db 1169 AlaAlaThrThrProProLysProSerTrpArgValGlnArgSerGlnAsnProLeu 1188
QY 3385 AAGCTTGGCTCCATCGGTGGTTCTT 3411
Db 1189 LysIleArgLeuThrArgGlyAlaPro 1197
RESULT 8
ADH76470
ID ADH76470 standard; protein; 1199 AA.
XX
AC ADH76470;
XX
DT 15-APR-2004 (first entry)
XX Murine leukemia virus pol gene derived protein.
DE
XX chimeric plasmid; replicative retroviral genome; gag; pol;
KW murine leukemia virus; MLV; env; gibbon ape leukemia virus; GalV; virion;
KW MLV-GalV-type; gene therapy; pAM.
XX Murine leukemia virus.
OS
XX

PN FR2832424-A1.
XX 23-MAY-2003.
XX 20-NOV-2001; 2001FR-00014976.
XX PF 20-NOV-2001; 2001FR-00014976.
XX PR (GENE-) GENETHON III.
XX PA Audit M, Cosset FL;
XX FI WPI; 2003-471779/45.
XX DR Chimeric plasmid containing replicative retroviral genome, useful for
XX PT making positive control virions in testing for replication-competent
XX PT retrovirus.
XX PS Claim 4; SEQ ID NO 2; 70pp; French.
XX CC The invention relates to a novel chimeric plasmid comprising a
XX CC replicative retroviral genome. The replicative retroviral genome
XX CC comprises: the gag and pol sequences from a murine leukemia virus (MLV);
XX CC and a chimeric env sequence comprising regions corresponding to parts of
XX CC the envelope derived from: an MLV genome; and a gibbon ape leukemia virus
XX CC (GALV). Virions produced by expressing the viral genome of the chimeric
XX CC plasmid are useful as positive controls in a test for detection of
XX CC replication-competent retroviruses in preparations of MLV-GALV-type
XX CC retroviral vectors. For example, to ensure that the MLV-GALV-type
XX CC retroviral vectors, intended for gene therapy, have no capacity for
XX CC replication. This sequence represents the protein derived from an MLV pol
XX CC gene from a pM plasmid of the invention.
XX SQ Sequence 1199 AA;

Alignment Scores:
Pred. No.: 0 Length: 1199
Score: 2808.50 Matches: 722
Percent Similarity: 76.2% Conservative: 160
Best Local Similarity: 62.3% Mismatches: 245
Query Match: 61.6% Indels: 31
DB: 7 Gaps: 11

US-10-723-552-3_COPY_2307_5741 (1-3435) x ADH76470 (1-1199)

QY 4 GGTGCCACAGGCAACACAGATATCCATGGACTACCCGAGAACAGTTGACTTGGAGTG 63
Db 51 GlyAlaThrGlyGlyLysArgTyrArgTyrThrThrAspArgLysValHisLeuAlaThr 70
QY 64 GGACGGGTAAACCACTCGTTCTGCTCATCTAGTGCACGACCCCTCTTAGGTAGA 123
Db 71 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 90
QY 124 GACTTATTGACCAAGATGGGACCAAAATTTCTTTTGA---CAAGGGAACACAGAGTG 180
Db 91 AspLeuLeuThrLysLeuLysAlaGlnLeHisPheGluGlySerGlyAlaGlnValMet 110
QY 181 TCTGCAATAAACAACCTATCATCTGTGTGACCCCTCAATTAGATGACGAATATCGACTA 240
Db 111 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAenlleGluAspGluHisArgLeu 130
QY 241 TACTCTCCCTAGTAAGCCTGATCAAAATATA---CAATCTGGTTGGACAGTTTCCC 297
Db 131 HisGluThrSerLysGluProAspValSerLeuGlySerThrTyrTrpLeuSerAspPhePro 150
QY 298 CAAGCCTGGGCAAAACCGAGGATGGTTGGCAAAAGCAAGTTCCCCCAACAGTTATT 357
Db 151 GlnAlaTyrAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuLeulle 170
QY 358 CAACCTAAGGCGCAGTCCACACAGTGTGACGTACAGACATCCCTTGTAGTAAGAAGCT 417
Db 171 ProLeuLysAlaThrSerThrProValSerlleLysGlnTyrProMetSerGlnGluAla 190

QY 418 CAAGAAGGAATTCGGCCGCATGTCCAAGAGTAATCAACAGGGCATCTAGTTCCTGTC 477
Db 191 ArgLeuGlylleLysProHisIleGlnArgLeuLeuAspGlnGlylleLeuValProCys 210
QY 478 CAATCTCCCTGGAATACTCCCTGCTACCGGTTAGAAAGCCTGGAGCTAATGACTATCGA 537
Db 211 GlnSerProTyrPheThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 230
QY 538 CCAGTACAGGACTTGAGAGAGGTCAATAAAGGGTGCAGGATATACACCAACAGTCCCG 597
Db 231 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 250
QY 598 AACCTTTATAACCTCTTGTGTCTCCACCCCAACCGAGCTCGTATACAGTATTGGAC 657
Db 251 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTyrTyrThrValLeuAsp 270
QY 658 TTAAGGATGCCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTGCCTTC 717
Db 271 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 290
QY 718 GAATGGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGACCCGACTGCCCAA 777
Db 291 GluTyrArgAspProGluMetGlyIleSerGlyGlnLeuThrTyrThrArgLeuProGln 310
QY 778 GGGTTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACTCGCCCAACTTC 837
Db 311 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 330
QY 838 AGGATCCAAACACCTCAGGTGACCTCTCCTCAGTACGTGGATGACCTGCTTCTGGCGGA 897
Db 331 ArgIleGlnHisProAspLeuLeuLeuGlnTyrValAspAspLeuLeuLeuAlaAla 350
QY 898 GCCACCAACAGGACTCTTAGAAGCAGCAGGACCTACTCTGCTGGAATCTCTGACCTA 957
Db 351 ThrSerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu 370
QY 958 GGCTACAGAGCTCTGCTAAGAAGGCCAGATTTTGGAGAGAGAGTAAACATCTTGGGG 1017
Db 371 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 390
QY 1018 TACAGTTTGGCGGACGGCAGCGATGCTGACGAGGACCGAAGAAACCTGTAGTCCAG 1077
Db 391 TyrLeuLeuLysGluGlyGlnArgTyrLeuThrGluAlaArgLysGluThrValMetGly 410
QY 1078 ATACGGCCCCCAACCAACAGCCAAACAAATGAGAGATTTTGGGACAGCTGGATTTCG 1137
Db 411 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 430
QY 1138 AGACTGTGGATCCCGGGTTTTCGACCTTAGCAGCCCCACCTCTACCCGCTAACCAAGAA 1197
Db 431 ArgLeuTyrPheProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 450
QY 1198 AAAGGGGAATTCCTCTGGGCTCTGAGCACCAAGGAGCATTTGATGCTATCAAAAGGCC 1257
Db 451 GlyThrLeuPheAsnTyrGlyProAspGlnGlnLysAlaTyrGlnGluIleLysGlnAla 470
QY 1258 CTGCTGACGCGACCTGCTCTGCGCCCTCCCTGACGTAACCTTTTACCTTTATCTG 1317
Db 471 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 490
QY 1318 GATGAGCTAAGGAGTAGCCCGGGAGTTTAAACCCCAACCCCTAGGACCATGAGAGAAGA 1377
Db 491 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTyrArgArg 510
QY 1378 CCTGTGCTGCTACCTGTCAAAGAAGCTCGATCCTGTAGCCAGTGGTGGCCCATATGCGCTG 1437
Db 511 ProValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTyrProProCysLeu 530
QY 1438 AAGGCTATCGAGCTGGCCATACTGCTAAGGACGCTGACAACTTCTGCTTGGGACAG 1497
Db 531 ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGln 550
QY 1498 AATAAATCTGTAATAGCCCCCATGATTTGGAGAACATCTGTTGGCAGCCCCCAGACCGA 1557

PN W0200042199-A1.
 PD 20-JUL-2000.
 XX 14-JAN-2000; 2000WO-US000896.
 XX 15-JAN-1999; 99US-0116099P.
 XX (MOLE-) MOLECULAR BIOLOGY RESOURCES.
 PA Swaminathan N;
 XX WPI; 2000-482830/42.
 DR
 XX Novel genes encoding reverse transcriptase polypeptides modified by
 PT altering or adding the integrase domains by truncation internally and/or
 PT at the C-termini, useful in cDNA synthesis and amplification procedures.
 XX
 PS Claim 1; Page 154-159; 189pp; English.
 XX This invention relates to isolated polynucleotide sequences encoding a
 CC polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT)
 CC activity. RTs are found in a variety of retroviruses and their defining
 CC activity is the ability to synthesise a cDNA strand using an RNA
 CC template. The invention includes sequences AA87808-A87822 and AA87840-
 CC A87842 which represent nucleotide sequences encoding RT polypeptides.
 CC Included in the invention are reverse transcriptase protein sequences
 CC AA81298-B12995. Sequences AA87829-A87839 and AA87843-A87897 represent
 CC oligonucleotides used in the identification and synthesis of the RT
 CC nucleotide sequences of the invention. Many of the RT nucleotide
 CC sequences encode modified RT proteins, which exhibit improved stability
 CC and/or improved solubility, relative to naturally occurring reverse
 CC transcriptases. The modified RT nucleotide and protein sequence are used
 CC in improved polymerase chain reaction (PCR) methods. The nucleotide
 CC sequences can be used in sequencing methods. The present sequence
 CC represents a reverse transcriptase protein of the invention
 XX
 SQ Sequence 1199 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 1199
 Score: 3797.50 Matches: 719
 Percent Similarity: 76.3% Conservative: 165
 Best Local Similarity: 62.0% Mismatches: 242
 Query Match: 61.4% Indels: 33
 DB: 3 Gaps: 12
 US-10-723-552-3_COPY_2307_5741 (1-3435) x AA812994 (1-1199)
 QY 4 GGTGCCACGGGCAACACAGTATCCATGGACTACCCGAGCAACAGTTGACTGGAGTG 63
 DB 51 GlyAlaThrGlyGlyLysArgGlyArgGlyrPrThrAspArgLysValHisLeuAlaThr 70
 QY 64 GGACGGTAACCCACTCGTTCTGTGTATCATCTGAGTCCGACACCCCTCTTAGGTAGA 123
 DB 71 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 90
 QY 124 GACTTATTGACACAGATGGGACGACAAATTTCTTTTGA---CAAGGGAACCCAGAGTG 180
 DB 91 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValMet 110
 QY 181 TCTGCAATAACAACCTATCACTGTGTGTGACCTCCCAATTAGATGACGAATATCGACTA 240
 DB 111 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAenIleGluAspGluHisArgLeu 130
 QY 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATA---CAATTCTGTGTTGGAACAGTTTCCC 297
 DB 131 HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 150
 QY 298 CAAGCTGGGACGAAACCCGACGGATGGTTGGCAAGCAAGTTCCTCCCAAGATTATT 357
 DB 151 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIle 170

QY 358 CAACTGAAGCCAGTGCACACACAGTGTCTCAGTCAGACAGTACCCCTTTGAGTAAAGACT 417
 DB 171 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrProMetSerGlnGluAla 190
 QY 418 CAAGAAGGAATTCGGCGCATGTCCAAAGATTAAATCAACAGAGGCATCTTAGTTCTCTGTC 477
 DB 191 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 210
 QY 478 CAATCTCCCTGGAATACTCCCTGCTACCGTTAGAAAGCCTGGGACTATGACTATCGA 537
 DB 211 GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 230
 QY 538 CCAGTACAGGACTTGAGAGAGGTCAATAAAGCGGTGCAGGATATACACCCCAACAGTCCCG 597
 DB 231 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 250
 QY 598 AACCTTTAATACCTCTTTGTGTCTCCACCCCAACGAGCTGTGTATACAGTATTGGAC 657
 DB 251 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp 270
 QY 658 TTAAGGATGCCTTCTCTGCTGAGATTACCCCACTAGCCCAACCACTTTTTCCTTC 717
 DB 271 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 290
 QY 718 GAATGGAGAGATCCAGGTACCGGGAAGAACCGGGCAGCTCACCTGAGACCCGACTGCCCAA 777
 DB 291 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 310
 QY 778 GGGTTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTC 837
 DB 311 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 330
 QY 838 AGGATCCAAACCTCAGGTGACCTCTCAGTACCTGATGATGACCTGCTCTCGCGGGA 897
 DB 331 ArgIleGlnHisProAspLeuIleLeuLeuGlnTyrValAspAspLeuLeuAlaAla 350
 QY 898 GCACACCAACAGGACTGCTTAGAAGCAGCAAGGACACTACTGTGGAATTCTCTGACCTA 957
 DB 351 ThrSerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu 370
 QY 958 GGCTACAGAGCTCTGCTAAGAAGCCGACAGATTTCGAGGAGAGAGAGTAACATCTGGGG 1017
 DB 371 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 390
 QY 1018 TACAGTTTGGGACGGGACCGATGCTGACGGGACGACGAGAAACCTAGTCCAG 1077
 DB 391 TyrLeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly 410
 QY 1078 ATACCGGCCCAACACACAGCCCAAAATGAGAGAGTTTTTTGGGACAGCTGGATTTCG 1137
 DB 411 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 430
 QY 1138 AGACTGTGATPCCCGGGTTTGGACCTTACAGACCCCACTTACCCGCTAACCAAGAA 1197
 DB 431 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 450
 QY 1198 AAAGGGGAATCTCTCGGCTCTGAGCACACAGAGGCAATTTGATGCTATCAAAAGGCC 1257
 DB 451 GlyThrLeuPheAsnTrpGlyProAspGlnGlnLysAlaTyrGlnGluIleLysGlnAla 470
 QY 1258 CTGCTGAGCGCACCTGCTGCGCCCTCCTGACCTAACTAAACCTTTTACCTTTATGTG 1317
 DB 471 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 490
 QY 1318 GATCAGCGTAAGGAGTAGCCCGGGAGTTTAAACCAACCCCTAGGACCATGAGAGA 1377
 DB 491 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArgArg 510
 QY 1378 CCTCTCGCTACCTCTCAAGAAGCTGATCCTCTAGCCACTGCTGGTGGCCCATATGCTGT 1437
 DB 511 ProValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTrpProProCysLeu 530
 QY 1438 AAGGCTATCGCAGCTGTGGCCATCTGCTCAAGGACGCTGACAAATTGACTTTTGGGACAG 1497

```
Db 531 ArgMetValAlaAlaAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln 550
Qy 1498 AATAAATAGCCATCATCGTGGAGAACATCGTTGGCAGCCCGACGACGA 1557
Db 551 ProLeuValIleLeuAlaProHisAlaValGluAlaLeuValLysGlnProAspArg 570
Qy 1558 TGGATGACCAACGCCCGCATGACCCACTATCAAGCCCTGCTCTC---ACAGAGAGGTC 1614
Db 571 TrpLeuSerAsnAlaArgMetHisTyrglnAlaLeuLeuLeuLeuAspThrAspArgVal 590
Qy 1615 ACGTTCGCTCCACAGCGCTCTCAACCTGCCACTCTTCCTCGCTGAAGAGACTGATGAA 1674
Db 591 GlnPheGlyProValAlaLeuAsnProAlaThrLeuLeuPro---LeuProGluGlu 609
Qy 1675 CCAGTGCATCATGATTGGCCATCACTATTGATTGAGGAGACTGGGGTCCGCAAGACCTT 1734
Db 610 GlyLeuGlnHisAsnCysLeuAspIleLeuAlaGluAlaHisGlyThrArgProAspLeu 629
Qy 1735 ACAGACATACCGCTGACTGGAGAAAGTGTAACTGTTCTACTGACGGAGCAGCTATGTG 1794
Db 630 ThrAspGlnProLeuProAspAlaAspHisThrTrpTyThrAspGlySerSerLeuLeu 649
Qy 1795 GTGNAAGTAAAGATGCTGGCGCGCGTGTGGACGGGACCCGACCATCTGGGCC 1854
Db 650 GlnGluGlyGlnArgLysAlaGlyAlaAlaValThrThrGluThrGluValIleTrpAla 669
Qy 1855 AGCAGCTGCCGGAAGAACTTCAGCACAAAGCTGAGTCATGCGCCCTCACCAAGCT 1914
Db 670 LysAlaLeuProAlaGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAla 689
Qy 1915 TTGGCGCTGGCGAAGGAAATCCATAACATTTATACGACAGCAGGTATGCTTTGCG 1974
Db 690 LeuLysMetAlaGluGlyLysLysLeuAsnValThrThrAspSerArgTyAlaPheAla 709
Qy 1975 ACTGCACAGTACATGGGGCATCTATAACAAGGGGTTGCTTACTCTACGAGGGAGG 2034
Db 710 ThrAlaHisIleHisGlyGlyIleTyArgArgArgGlyLeuLeuThrSerGluGlyLys 729
Qy 2035 GAAATAAGAACAAAGAGAAATCTAAGCGCTATTAGAACCGTACATTTTACCAAAAGG 2094
Db 730 GluIleLysAsnLysAspGluIleLeuAlaLeuLysAlaLeuPheLeuProLysArg 749
Qy 2095 CTAGCTATTATACACTGCTCGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAC 2154
Db 750 LeuSerIleIleHisCysProGlyHisGlnLysGlyHisSerAlaGluAlaArgGlyAsn 769
Qy 2155 CAGATGCTGACCGGTTGCCAAGCAGGACCCAGGCTTTAACTTCTGCTCTATATA 2214
Db 770 ArgMetAlaAspGlnAlaAlaArgLysAlaAla-----IleThr 782
Qy 2215 GAAATGCCAAAGCC-----CCAGAACCCAGA 2241
Db 783 GluThrProAspThrSerThrLeuLeuIleGluAsnSerSerProTyThrSerGluHis 802
Qy 2242 CGACAGTACACCTAGAGACTGCGCAAGAGATAAAGAGATAGACACAG---TTCCTGAG 2298
Db 803 PheHisTyThrValThrAspIleLysAspLeuThrLysLeuGlyAlaIleTyAspLys 822
Qy 2299 ACTCCGAGGAGGACTGCTATACCTCAGATGGGAAGAAATCCCTGCCCCCAAGAGGG 2358
Db 823 ThrLysLysTyTrpValTyLys-----GlnGlyLysProValMetProAspGlnPheThr 840
Qy 2359 TTAGAATATGTCACAGATACATCGTCTAACCCACCTAGGAACCTAAACCTGCAGCAG 2418
Db 841 PheGluLeuLeuAspPheLeuHisGlnLeuThrHisLeuSerPheSerLysMetLysAla 860
Qy 2419 TTGCTCAGAAACA-----TCCCTTATCATGCTTCTGAGGCTACCCAGGAGTGGCTGACTCG 2472
Db 861 LeuLeuGluArgSerHisSerProTyTyTrpMetLeuAsnArgAspArgThrLeuLysAsn 880
Qy 2473 GTGTCACAAATGTGTGGCCCTGCCAGCTGGTTAATCTTAATCTTCCAGAAATGCCCTCCA 2532
Db 881 IleThrGluThrCysLysAlaCysAlaGlnValAsnAlaSerLysSerAlaValLysGln 900
Qy 2533 GGGAGAGACTAAGGGGAGCCACCCAGCGCTCCTCGGAGTGGAGTCTCAGTACGAGTA 2592
Db 901 GlyThrArgValArgGlyHisArgProGlyThrHisTrpGluIleAspPheThrGluIle 920
Qy 2593 AAGCCGGCTAAATACGGAACCAATACTATTGGTTTTTTGTAGACACCTTTTCAGATGG 2652
Db 921 LysProGlyLeuTyGlyTyLysTyLeuLeuValPheIleAspThrPheSerGlyTrp 940
Qy 2653 GTAGAGCTTATCTACTAAGAAAGAGACTTCAACCGTGGTGGCTGAAAAAATACTGAA 2712
Db 941 IleGluAlaPheProThrLysLysGluThrAlaLysValValThrLysLysLeuGlu 960
Qy 2713 GAAATTTTCCAAATTTGGAAATACCTAAGGTAAATAGGTTCAGACAAATGGTCCACTTTT 2772
Db 961 GluIlePheProArgPheGlyMetProGlnValLeuGlyThrAspAsnGlyProAlaPhe 980
Qy 2773 GTTCCCGAGGTAACTCAGGACTCGCCCAAGATATTGGGGATTGATTGGAAACTGCTATTGT 2832
Db 981 ValSerLysValSerGlnThrValAlaAspLeuLeuGlyIleAspTrpLysLeuHisCys 1000
Qy 2833 GCATACAGACCCCAAGACTCAGACAGTAGAGAGATGAATGAACCAATTAAGAGACC 2892
Db 1001 AlaTyArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThr 1020
Qy 2893 CTTACTAAATTCACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCCTGCGCTTTGTG 2952
Db 1021 LeuThrLysLeuThrLeuAlaThrGlySerArgAspTrpValLeuLeuLeuProLeuAla 1040
Qy 2953 CTTTATAGGTTTAGGAACACCCCTGGAGCAGTTTGGGCTGACCCCTATGAATTAATCTTAC 3012
Db 1041 LeuTyArgAlaArgAsnThrProGlyProHisGlyLeuThrProTyGluIleLeuTy 1060
Qy 3013 GGGGGACCCCGCCCATTTGGTAGAAATGCTTCTCTACATAGTGTGCTGCTGCTGCTTTC 3072
Db 1061 GlyAlaProProProLeuValAsnPheProAspProAspMetThrArgValThrAsnSer 1080
Qy 3073 CAGCTTTGTTCTCTAGCTCAAGGCACTTGAGGTGTGAGACACGAGCGTGGAGGCAA 3132
Db 1081 ProSerLeuGlnAlaHisLeuGlnAlaLeuTyLeuValGlnHisGluValTrpArgPro 1100
Qy 3133 CTCGGAGGCTCTACTCAGGAGGAGGACTTGCAG---ATCCACATCTGTTTCCAAAGTG 3189
Db 1101 LeuAlaAlaIleTyGlnGluGlnLeuAspArgProValValProHisProTyArgVal 1120
Qy 3190 GGAGATTACGTCTACTAGTACCGCCCGTGCAGGAAACCTCGAGACTCGGTGGAAGGGC 3249
Db 1121 GlyAspThrValTrpValArgHisGlnThrLysAsnLeuGluProArgTrpLysGly 1140
Qy 3250 CCTTATCTCTCTCTTTTGCACACCAACCGCTGTGAAGTCCGAAGAAATCTCCACCTGG 3309
Db 1141 ProTyThrValLeuLeuThrProThrAlaLeuLysValAspGlyIleAlaAlaTrp 1160
Qy 3310 ATCCATGCTATCCACGTTAAACCGCGCCACCT-----CCCGATTG-----GGG 3354
Db 1161 IleHisAlaAlaHisValLysAlaAlaAspProGlyGlyGlyProSerSerArgLeuThr 1180
Qy 3355 TGGAAAGCCGAAAGACTGAAATCCCTTAAAGTTCGCTCCATCCGCTGCTTCTCT 3411
Db 1181 TrpArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGluAlaPro 1199
```

RESULT 10

AAB10044

ID AAB10044 standard; protein; 1737 AA.

XX AAB10044;

XX 02-NOV-2000 (first entry)

XX MMLV gag-pol protein.

XX Glycoprotein; gag gene; pol gene; GP-1; GP-2; anti-HIV; cytostatic;

KW gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;
KW Carcinoma; melanoma.
XX Moloney murine leukemia virus.
XX EP1006196-A2.
XX PD 07-JUN-2000.
XX 25-NOV-1999; 99EP-00250415.
XX PR 26-NOV-1998; 98DE-01056463.
XX (PETT-) PETTE INST HEINRICH.
XX Von Laer MD;
FI WPI; 2000-378268/33.
DR New retroviral packing cell useful as pharmaceutical carrier in gene
PT therapy for treatment of HIV and neoplasms, comprises retroviral genes
PT and glycoproteins.
XX Disclosure; Page 32-36; 69pp; German.
XX This invention describes a novel retroviral packing cell (I), comprising
CC the retroviral genes gag, pol and glycoproteins GP-1 and GP-2 of the LCMV
CC coding gene gp, or a part of these. The products of the invention have
CC anti-HIV and cytostatic activity and can be used for gene therapy. (I) is
CC useful for in vitro infection of cells, especially hematopoietic stem
CC cells, for expression of transgenes in cells and as a pharmaceutical
CC carrier for gene therapy. (I) is therefore useful in the treatment of
CC infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and
CC other diseases. This sequence represents the Moloney murine leukemia
CC virus (MMuLV) gag-pol protein which is described in the method of the
CC invention
XX
SQ Sequence 1737 AA;

Alignment Scores:
Pred. No.: 0 Length: 1737
Score: 3797.50 Matches: 719
Percent Similarity: 76.3% Conservative: 165
Best Local Similarity: 62.0% Mismatches: 242
Query Match: 61.4% Indels: 33
DB: 3 Gaps: 12

US-10-723-552-3_COPY_2307_5741 (1-3435) x AAB10044 (1-1737)

QY 4 GGTGCCAGGGCAACACAGTATCCATGGACTACCGAGACAGTGTGACTTGGAGTG 63
Db GlyAlaThrGlyGlyLysArgTyrArgTrpThrThrAspArgGlyValHisLeuAlaThr 608

QY 64 GGACGGGTAAACCACTCTGCTGTCATCTGAGTGCACCGAGCCACCTCTTAGGTAGA 123
Db GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProleuLeuGlyArg 628

QY 124 GACTTATTGACAGATGGGACCAAAATTTCTTTTGAA---CAAGGAAACCAAGAGTG 180
Db AspLeuLeuThrLysLeuLysAlaGlnHisPheGluGlySerGlyAlaGlnValMet 648

QY 181 TCTGCAATACAAACCTATCACTGTGTGACCTCCATTTAGATGAGCAATATCGACTA 240
Db GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu 668

QY 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATA---CAATTCTGGTGTGAAACAGTTTCCC 297
Db HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 688

QY 298 CAAGCTGGGCGAGAAACCGAGGATGGTTGGCAAGCAAGTTTCCCCCAAGATTATT 357
Db GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 708

QY 358 CAACTGAAGGCCAGTGGCCACACACAGTGTCACTCAGACAGTACCCTCTGAGTAAAGAAGCT 417
Db ProLeuLysAlaThrSerThrProValSerIleLysGlnIlePheSerGlnGluAla 728

QY 418 CAAGAAGGAATTCGGCCGCATGTCCTCAAGATTAATCCAACAGGGCATCTCTAGTTCTCTGTC 477
Db ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 748

QY 478 CAATCTCCCTGGAAATACCTCCCTGCTACCGGTTAGAAAGCCTGGAGCAATAAGTACTATCGA 537
Db GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 768

QY 538 CCAGTACAGGACTTTCAGAGAGTCAATAAAGCGGTGCAGGATATACACCCACAGTCCCG 597
Db ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 788

QY 598 AACCTTTATACCTCTTGTGTGCTCTCCACCCCAACCGAGCTGGTATACAGTATTGGAC 657
Db AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp 808

QY 658 TTAAGGATGCTCTTCTGCTGAGATTACACCCCTAGCCCAACCACTTTTTCCTTC 717
Db LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 828

QY 718 GAATGGAGAGATCAGGTACGGGAAGAACCGGGCAGCTCACCTGGACCCGACTGCCCAA 777
Db GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 848

QY 778 GGGTTCGAAGAACTCCCGACCATCTTTTCAGAAAGCCCTTACACAGAGACTCGGCCAATTC 837
Db GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 868

QY 838 AGGATCCAAACCTCAGGTGACCTCTCTCAGTACGTGGATGACCTGCTTCTGGCGGA 897
Db ArgIleGlnHisProAspLeuIleLeuLeuGlnIleValAspAspLeuLeuAlaAla 888

QY 898 GCCACCAACAGGACTGCTTAGAAGGACGACGAGCTACTCTGGAATTCCTGACCTA 957
Db ThrSerGluLeuAspCysGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu 908

QY 958 GGCTACAGAGCTCTGCTAAGAAGGCCAGATTTTCAGGAGAGAGGTAACTACTTGGGG 1017
Db GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 928

QY 1018 TACAGTTTGGGAGCGGACGAGTGTGCTGACGAGCAGCGAGAAACCTGTAGTCCAG 1077
Db TyrLeuLysLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly 948

QY 1078 ATACGGGCCCAACACACAGCCAAACAAATGAGAGAGTGTTCGGGACAGCTGGATTTCG 1137
Db GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 968

QY 1138 AGACTGTGATCCCGGGTTTCGACCTTTAGCAGCCCTTACAGCCCTTACCCGCTAACCAAGAA 1197
Db ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 988

QY 1198 AAAGGGGAATTCCTCTGGGCTCTGAGCACACAGAGGCAATTTGATGCTATCAAAAGGCC 1257
Db GlyThrLeuPheAsnTrpGlyProAspGlnLysAlaTyrGlnGluIleLysGlnAla 1008

QY 1258 CTGCTGAGCGCACCTGCTCTGGCCCTCCCTCAGCTAACTAAACCTTTTACCTTTATCTG 1317
Db LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 1028

QY 1318 GATGAGCGTAAGGAGTAGCCCGGGAGTTCCTTAAACCCAAACCTTAGGACCATGGAGAGA 1377
Db AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArgArg 1048

QY 1378 CCTGCTGCTACCTGTCGAAGAGCTGATCCTCTAGCCAGTGGTGGCCCATATGCTGT 1437
Db ProValAlaTyrLeuSerLysLysLeuAspProValAlaAlaGlyTrpProProCysLeu 1068

QY 1438 AAGGCTATCGCAGCTGTGGCCCATCTGCTCAAGGACGCTGACAAATTCGACTTTGGGACAG 1497

KW murine VL30; retroviral vector; biotechnology; pharmaceutical;
 KW gene therapy; Moloney murine leukemia virus; pol gene.
 XX Moloney murine leukemia virus.
 XX WO9925862-A2.
 XX 27-MAY-1999.
 XX 19-NOV-1998; 98WO-US024667.
 XX 19-NOV-1997; 97US-0066148P.
 XX (NATU-) NATURE TECHNOLOGY INC.
 XX Hodgson CP, Zink MA, Guoping X;
 DR WPI; 1999-347491/29.
 DR N-PSDB; AAX77178.
 XX New chimeric packaging signal useful in gene therapy.
 PT Disclosure: Fig 4; 6lpp; English.
 PS The invention relates to a new chimeric viral packaging signal that
 CC comprises: (i) essential packaging nucleic acid sequence, from a
 CC mammalian type C retrovirus, functionally joined to (ii). (ii) at least
 CC one non-essential packaging nucleic acid sequence (lacking the gag gene)
 CC from murine VL30. The chimeric viral packaging signals are used to
 CC produce retroviral vectors for transfer of genetic material in gene or
 CC cell therapy, or other biotechnological applications. Typical uses are in
 CC production of pharmaceuticals (e.g. enzymes, antibiotics, antibodies,
 CC antisense RNA, cytokines etc.) in animals or cell cultures. Elimination
 CC of the gag gene results in safer vectors (reduced formation of
 CC replication competent retrovirus as a result of recombination events
 CC involving the gag gene). Also vectors that contain the chimeric packaging
 CC signal have high infectious titers (about 1 million transduction units
 CC (TU)/ml), comparable with that for gag-containing vectors and higher than
 CC for known gag-free vectors. The present sequence represents the pol gene
 CC product of the MoMLV virus
 XX Sequence 1224 AA;

Alignment Scores:
 Pred. No.: 0 Length: 1224
 Score: 3781.00 Matches: 717
 Percent Similarity: 74.5% Conservative: 165
 Best Local Similarity: 60.6% Mismatches: 244
 Query Match: 61.2% Indels: 58
 DB: 2 Gaps: 11

US-10-723-552-3_COPY_2307_5741 (1-3435) x AAY17947 (1-1224)

QY 4 GGTGCCACAGGCAACACAGTATCCATGACCTACCGAAGAACAGTTGACCTGGAGTG 63
 Db 51 GlyAlaThrGlyGlyAspGlyArgTrpThrAspArgGlyValHisLeuAlaThr 70
 QY 64 GCACGGGTAAACCACTCGTTTCTGTCATACCTGAGTCCGACACCCCTCTTAGGTAGA 123
 Db 71 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 90
 QY 124 GACTTATTGACCAAGATGGGACCAAAATTTCTTTTGA---CAAGGGAAACAGAACTG 180
 Db 91 AspLeuLeuThrLysLeuLysAlaGlnHisPheGlyLysGlyValAlaGlnValMet 110
 QY 181 TCTCAAAATAACAAACCTACTCTGTGTGACCTCCAAATAGATGACGAATATCGACTA 240
 Db 111 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu 130
 QY 241 TACTCTCCCTAGTAAAGCTGATCAAAATATA---CAATTCTGTTTGAACAGTTTCCC 297
 Db 131 HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 150

QY 298 CAAGCTGGGCGAGAAACCGCAGGATGGGTTTGGCAAGCAAGTTCCCCCAAGATTATT 357
 Db 151 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleile 170
 QY 358 CAACTGAAGCCAGTGCACACAGTGTCTAGTACAGAGTACCCCTGAGTAAAGAGCT 417
 Db 171 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrProMetSerGlnGluAla 190
 QY 418 CAAGAAGGATTCGGCGCATGTCCAAGATTATCCAAACAGGCGATCTAGTTCCCTGTC 477
 Db 191 ArgLeuGlyLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 210
 QY 478 CAATCTCCCTGGAATACTCCCTCTACCGGTTAGAAAGCCTGGGACTTAATGACTATCGA 537
 Db 211 GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyrArg 230
 QY 538 CCAGTACAGGACTTGAGAGAGTCAATAACGGGTGCGAGGATATACACCCCAAGTCCCG 597
 Db 231 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 250
 QY 598 AACCTTTATACCTCTTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGAC 657
 Db 251 AsnProTyrAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyrThrValLeuAsp 270
 QY 658 TTAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTGCCTTC 717
 Db 271 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 290
 QY 718 GAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGTCTACCTCGAACCCGACTGCCCAA 777
 Db 291 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 310
 QY 778 GGGTTCAGAACTCCCGACCATCTTTCACGAAGCCCTACACAGAGACCTGGCCCACTTC 837
 Db 311 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 330
 QY 838 AGGATCCAAACACCTCAGGTGACCCCTCTCCAGTACGTGAGTACCTGCTTCTGGCGGGA 897
 Db 331 ArgIleGlnHisProAspLeuIleLeuLeuGlnTyrValAspAspLeuLeuLeuAlaAla 350
 QY 898 GCCACCAACAGGACTGCTTAGAAGGCAACGAGGACCTACTGTGGAAATTTCTGACCTA 957
 Db 351 ThrSerGluLeuAspCysGlnGlnGlyThrArgAlaLeuLeuGlnThrLeuGlyAsnLeu 370
 QY 958 GGCTACAGAGCTCTGCTTAGAAGCCAGATTTTCAGGAGAGAGGTAACTACTTGGGG 1017
 Db 371 GlyTyrArgAlaSerAlaLysLysAlaGlnIleCysGlnLysGlnValLysTyrLeuGly 390
 QY 1018 TACAGTTTGGCGGACGGGCGAGTGGCTGACGGAGGCACCGAAGAAACCTGTAGTCCAG 1077
 Db 391 TyrLeuLeuLysGluGlyGlnArgTrpLeuThrGluAlaArgLysGluThrValMetGly 410
 QY 1078 ATACCGGCCCCAACACAGCCCAACAAATGAGAGAGTGTGTTTGGGACAGCTGGATTGTC 1137
 Db 411 GlnProThrProLysThrProArgGlnLeuArgGluPheLeuGlyThrAlaGlyPheCys 430
 QY 1138 AGACTGTGGATCCCGGGTTGGACCTTACGACCCCACTCTACCCGCTACCAACAAAGAA 1197
 Db 431 ArgLeuTrpIleProGlyPheAlaGluMetAlaAlaProLeuTyrProLeuThrLysThr 450
 QY 1198 AAAGGGGAATTCCTCTGGGCTCTGAGCACCAGGACCATTTGATGCTATCAAAAAGGCC 1257
 Db 451 GlyThrLeuPheAsnTrpGlyProAspGlnGlnLysAlaTyrGlnGluLysGlnAla 470
 QY 1258 CTGTGACGCGCACTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTTACCCTTTATGTG 1317
 Db 471 LeuLeuThrAlaProAlaLeuGlyLeuProAspLeuThrLysProPheGluLeuPheVal 490
 QY 1318 GATGAGCGTAGGAGTAGCCCGGGGAGTTTAAACCCCAACCCCTAGGACCATGAGAGA 1377
 Db 491 AspGluLysGlnGlyTyrAlaLysGlyValLeuThrGlnLysLeuGlyProTrpArgArg 510
 QY 1378 CCTGTGCGCTACTGTCAAAGAAGCTCGATCCTCTGTAGCCAGTGGTGGCCCATATGCTG 1437

Db	511	ProValAlaIatyrLeuSerLysLeuAspProValAlaIaGlyTrpProProCysLeu	530
Qy	1438	AAGCGTATCGCAGCTGTGGCCATATCTGCTCAAGGACGCTGACAAATTCGACTTTGGGACAG	1497
Db	531	ArgMetValAlaAlaIleAlaValLeuThrLysAspAlaGlyLysLeuThrMetGlyGln	550
Qy	1498	AATATACTGTAATAGCCCCCATGCAATGTGGAGAACATCGTTGGCAGCCCCCAGACCGA	1557
Db	551	ProLeuValIleLeuAlaProHisAlaValGluAlaLeuValLysGlnProProAspArg	570
Qy	1558	TGGATGACCAAGCCCGCATGACCCACTATCAAGCCCTGCTCTC---	1614
Db	571	TrpLeuSerAenAlaArgMetThrHisTyrGlnAlaLeuLeuAspThrAspArgVal	590
Qy	1615	ACGTTGCGTCCACACGCGCTCTCAACCTCGCACTTCTGCTGGAAGAGACTGATGAA	1674
Db	591	GlnPheGlyProValValAlaLeuAsnProAlaThrLeuLeuPro---	609
Qy	1675	CCAGTGACTCATGATTGCCATCAACTATTTGATTGAGGAGACTGGGGTCCCGAAGACCTT	1734
Db	610	GlyLeuGlnHisAenCysLeuAspIleLeuAlaGluAlaHisGlyThrArgProAspLeu	629
Qy	1735	ACAGACATACCGCTGACTGGGAGAAGTGTCTAACTGGTTTCACTGACGGAAGCAGCTATGTG	1794
Db	630	ThrAspGlnProLeuProAspAlaAspHisThrTrpTyrThrAspGlySerSerLeuLeu	649
Qy	1795	GTGGAAGTAAAGAGATGGCTGGGGCGCGTGTGGACGGGCCCGCAGCATCTGGGCC	1854
Db	650	GlnGluGlyGlnArgLysAlaGlyAlaAlaValThrThrGluThrGluValIleTrpAla	669
Qy	1855	AGCAGCTTCGCGAAGGAACCTTCAGCACAAAGGCTGAGCTCATGGCCCTCACCAAGCT	1914
Db	670	LysAlaLeuProAlaGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGlnAla	689
Qy	1915	TTGGCGCTGGCGAAGGGAATCCATAAACATTTATACGGACAGCAGGTATGCTTTGCG	1974
Db	690	LeuLysMetAlaGluGlyLysLysLeuAenValTyrThrAspSerArgTyrAlaPheAla	709
Qy	1975	ACTGCACATGATGGGGCCATCTATAACAAAGGGGTGTCTTACTCAGCAGGGCAGG	2034
Db	710	ThrAlaHisIleHisGlyGluIleTyrArgArgArgGlyLeuLeuThrSerGluGlyLys	729
Qy	2035	GAATAAAGAACAAAGAGGAATCTTAAGCCTATTAGAAGCGGTACATTTACCAAAAGG	2094
Db	730	GluIleLysAenLysAspGluIleLeuAlaLeuLeuLysAlaLeuPheLeuProLysArg	749
Qy	2095	CTAGCTATTATACACTGCTCGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAC	2154
Db	750	LeuSerIleIleHisCysProGlyHisGlnLysGlyHisSerAlaGluAlaArgGlyAen	769
Qy	2155	CAGATGCTGACCGGTGGCCACAGCGCAGCCCGAGGTGTAACTTCTGCTCTATAATA	2214
Db	770	ArgMetAlaAspGlnAlaAlaArgLysAlaAla-----IleThr	782
Qy	2215	GAATGCCCCAAAGCC-----CCAGAACCCAGA	2241
Db	783	GluThrProAspThrSerThrLeuLeuIleGluAenSerSerProTyrThrSerGluHis	802
Qy	2242	CGACAGTACACCTTAGAAGACTGGCAGAGATAAAAAAGATAGACACAG-----TTCTCTGAG	2298
Db	803	PheHisTyrThrValThrAspIleLysAspLeuThrLysLeuGlyAlaIleTyrAspLys	822
Qy	2299	ACTCCGGAGGGACCTGCTATACCTCAGATGGGAGGAATCTGCGCCCCCAAGAGGG	2358
Db	823	ThrLysLysTyrTrpValTyr-----GlnGlyLysProValMetProAspGlnPheThr	840
Qy	2359	TTGAATAATGTCCAAACAGATACATCGTCTAACCCACCTAGGAACATAAACCTGCACAG	2418
Db	841	PheGluLeuLeuAspPheLeuHisGlnLeuThrHisLeuSerPheSerLysMetLysAla	860
Qy	2419	TTGTGTCAAGACA-----TCCCTTTATCATGTTCTTGAGGCTACAGAGGTGGCTGACTCG	2472

QY 1201 GGGGAATTCCTCGGCTCCTGAGCACAGAGGAGCATTTGATGCTATCAAAAAGGCCCTG 1260
Db 1032 ThrLeuPheGlnTrpGlyThrGluGlnGlnLeuAlaPheGluAlaGlnLeuArgLysAlaLeu 1051
QY 1261 CTGAGCGCACCTGCTCTGGCCCTCCCTGACGCTAACTAAACCTTTACCTTTATGTGGAT 1320
Db 1052 LeuSerSerProAlaLeuGlyLeuProAspIleThrLysProPheGluLeuPheIleAsp 1071
QY 1321 GAGCGTAAGGAGTAGCCGGGGAGTTTAAACCCAAACCTTAGGACCATGGAGAAGACCT 1380
Db 1072 GluAenSerGlyPheAlaLysGlyValLeuValGlnLysLeuGlyProTrpLysArgPro 1091
QY 1381 GTCCCTACCTGCTCAAGAAGCTCGATCTGTAGCCAGTGTGGCCCATATGCTGAAG 1440
Db 1092 ValAlaTyrLeuSerLysLysLeuAspThrValAlaSerGlyTrpProCysLeuArg 1111
QY 1441 GCTATCGCAGCTGTGGCCATCTGCTCAAGGACGCTGACAATTTGACTTTGGGACAGAAT 1500
Db 1112 MetValAlaAlaIleAlaIleLeuValLysAspAlaGlyLysLeuThrLeuGlyGlnPro 1131
QY 1501 ATAAGTGAATAGCCCCCATGCAATTTGGAGAACATCGTTGGCAGCCCCCAGACCGATGG 1560
Db 1132 LeuThrIleLeuThrSerHisProValGluAlaLeuValArgGlnProProAsnLysTrp 1151
QY 1561 ATGACCAACGCCCGCATGACCCACTATCAAGCCCTGCTTC---ACAGAGAGGCTCAGC 1617
Db 1152 LeuSerAsnAlaArgMetThrHisTyrGlnAlaMetLeuLeuAspAlaGluArgValHis 1171
QY 1618 TTGCTCCACCCAGCGCTCTCAACCTGCGCACTCTCTGCTGAAGAGACTGATGAACCA 1677
Db 1172 PheGlyProThrValSerLeuAsnProAlaThrLeuLeuProLeuProSerGlyLysPro 1191
QY 1678 -----GTGACTCATGATGCGCACTATTGATGAGGAGCTGGGGTCCGCAAGGAC 1731
Db 1192 ProArgLeuSerProAsp-----LeuAlaGluThrMetAlaGlnThrAsp 1206
QY 1732 CTTACAGACATACCGCTGACTGGAGAGTGTAACTGCTGTTTCACTGACGAAAGCAGCTAT 1791
Db 1207 LeuThrAspGlnProLeuProAspAlaAspLeuThrTrpTyrThrAspGlySerSerPhe 1226
QY 1792 GTGTGGAAGTAAGAGATGGCTGGGGCGCGGTGTGGACGGACCCGCACGATCTGG 1851
Db 1227 IleArgAsnGlyGluArgLysAlaGlyAlaAlaValThrThrGluSerGluValIleTrp 1246
QY 1852 GCCAGCAGCCTGCGGAGGAATCTTACGACAAAGGCTGAGCTCATGGCCCTCACGCAA 1911
Db 1247 AlaAlaSerLeuProProGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGln 1266
QY 1912 GCTTTGGCGCTGGCGAAGGAAATCCATAACATTTATACGACAGCAGGTATGCTTTT 1971
Db 1267 AlaLeuLysMetAlaLysGlyLysLeuThrValTyrThrAspSerArgTyrAlaPhe 1286
QY 1972 GCGACTGCACGTCATAGGGCCCTATATAACAAAGGGGTGTCTACTCAGCAGGG 2031
Db 1287 AlathrAlaHisValHisGlyGluIleTyrArgArgArgGlyLeuLeuThrSerGluGly 1306
QY 2032 AGGGAATAAAGAACCAAGAGGAATCTTAAGCCTATTAGAAGCCGTACATTTTACCAAAA 2091
Db 1307 LysGluIleLysAsnLysAsnGluIleLeuAlaLeuLeuGluAlaLeuPheLeuProLys 1326
QY 2092 AGCGTAGCTATTATACCTGCTCTGGACATCAGAAAGCTAAAGATCTATATCCAGAGGA 2151
Db 1327 ArgLeuSerIleIleHisCysProGlyHisGlnLysGlyAspSerProGlnAlaLysGly 1346
QY 2152 AACCATGCTGACCGGGTTGCCAGCAGCGACCC-----CAGGGTGTTAAC 2199
Db 1347 AsnArgLeuAlaAspAspThrAlaLysLysAlaAlaThrGluThrGlnSerSerLeuThr 1366
QY 2200 CTTCTGCTCT-----ATAATAGAAATGCCCAAGCCCGACAGCCAGACAGTACACC 2253
Db 1367 IleLeuProThrGluLeuIleGluGlyProLys-----ArgProProTrpGluTyrAsp 1384
QY 2254 CTAGAAGACTGGCAAGATATAAAGATAGACCAGTTCTCTGAGACTCCGGAAGGGACC 2313

Db 1385 AspSerAspLeuAspLeuValGlnLysLeuGluAlaHisTyrGluProLysArgGlyThr 1404
QY 2314 TGCTATACCTCAGATGGAGGAATCCTGCCCCACAAAGAGGGTTAGATATGTCCAA 2373
Db 1405 ---TrpTyrArgGlyLysThrIleMetProGluLysTyrAlaLysGluLeuIleSer 1423
QY 2374 CAGATACATCGTCTAAACCCACCTAGGAACCTAAACACCTGCAGCAGTGTGGTCAGAACATCC 2433
Db 1424 HisLeuHisLysLeuThrHisLeuSerAlaArgLysMetLysThrLeuLeuGluArgGlu 1443
QY 2434 -----CCTTATCATGTTCTGAGGCTACACGAGTGCTGACTCGGTG 2475
Db 1444 GluThrGlyPheTyrLeuProAsnArgAspLeuHisLeuArgGlnValThrGluSer--- 1462
QY 2476 GTCAAACATTGTGTGCCCTGCCAGCTGTTTAAATGCTAATCTCTCCAGATGCTCCAGG 2535
Db 1463 -----CysArgAlaCysAlaGlnIleAsnAlaGlyLysIleLysPheGlyProAsp 1479
QY 2536 AAGAGACTAAGGGAGGCCACCCAGGCGCTCACTGGGAAGTGGACTTCACTAGGTAAAG 2595
Db 1480 ValArgAlaArgGlyArgArgProGlyThrHisTrpGluValAspPheThrGluIleLys 1499
QY 2596 CCGCTTAATACGGAACAAATACCTATTGTTTTTTGTAGACACCTTTTTCAGGATGGGT 2655
Db 1500 ProGlyMetTyrGlyTyrLysTyrLeuLeuValPheIleAspThrPheSerGlyTrpAla 1519
QY 2656 GAGCTTATCTACTAAGAAAGAGACTTCAACCGTGTGTGGCTAAAAAATCTCGGAAGAA 2715
Db 1520 GluAlaTyrProAlaLysHisGluThrAlaLysValValAlaLysLysLeuLeuGluGlu 1539
QY 2716 ATTTTCCAAAGATTTGGAATACCTAAGGTAAATAGGGTCAGACAAATGCTCAGCTTTTGT 2775
Db 1540 IlePheProArgTyrGlyIleProGlnValLeuGlySerAspAsnGlyProAlaPheIle 1559
QY 2776 GCCCAGTAAGTCAGGAGCTGGCCAAAGATATGGGATTCATTGGAATTCATTTGTCGA 2835
Db 1560 SerGlnValSerGlnSerValAlaThrLeuLeuGlyIleAsnTrpLysLeuHisCysAla 1579
QY 2836 TACAGACCCCAAGCTCAGGACAGTAGAGAGTGAATGAATGAACCATTAAGAGACCTT 2895
Db 1580 TyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgSerIleLysGluThrLeu 1599
QY 2896 ACTAAATTCACCGCGAGACTGGCGTTAATGATTGATAGTCTCTCTGCCCTTTGTGCTT 2955
Db 1600 ThrLysLeuThrLeuGluThrGlySerLysAspTrpValLeuLeuLeuProLeuValLeu 1619
QY 2956 TTTAGGTTAGGAACCCCTCGACAGATTTGGGCTGACCCCTATGAATTCATCTACGGG 3015
Db 1620 TyrArgValArgAsnThrProGlyProHisGlyLeuThrProPheGluIleLeuTyrGly 1639
QY 3016 GGAACCCCGCCCTTGGTAGAAATTCCTCTGTATAGTGTCTGACGTGCTGCTTCCAG 3075
Db 1640 AlaProProMetAlaHisPhePheAspThrAspIleSerThrPheAlaThrSerPro 1659
QY 3076 CTTTGTCTCTAGGCTCAAGCAGCTTGAGTGGCTGAGACCAACGAGCGTGGAGCAACTC 3135
Db 1660 ThrMetGlnAlaHisLeuArgAlaLeuGlnLeuValGlnGluLeuGlnArgProLeu 1679
QY 3136 CGGAGGCTTACTCAGGAGGAGGAGACTTGCAG---ATCCCATCATGTTTCCAAAGTGGGA 3192
Db 1680 AlaAlaAlaTyrArgGlyLysLeuGluThrProValValProHisProPheLysProGly 1699
QY 3193 GATTACGTCTAGCTAGCGCCACCGCTGCAGGAACCTCGAGACTCGGTGGAGGGCCCT 3252
Db 1700 AspSerValTrpValArgArgHisGlnThrLysAsnLeuGluProArgTrpLysGlyPro 1719
QY 3253 TATCTCGTACTTTTGACCAACACACCGCTGTGAAAGTCGAAGAAATCTCCACTGGATC 3312
Db 1720 HisIleValLeuLeuThrThrProThrAlaLeuLysValAspGlyValAlaAlaTrpIle 1739
QY 3313 CATGATCCCACTTAAACCGCGCCACCT----- 3342

Db 1012 LeuTrpIleProGlyPheAlaGluLeuAlaLaProLeuTyrProLeuThrArgProGly 1031
Qy 1201 GGGGAATTCTCTGGCTCCTGAGCACCAGAGGCATTGTGCTATCAAAAGGCCCTG 1260
Db 1032 ThrLeuPheGlnTrpGlyThrGluGlnGlnLeuAlaPheGlnLeuArgLysAlaLeu 1051
Qy 1261 CTGAGCGCACTGCTCTGGCCCTCCCTGACGTAACTAAACCTTTACCTTTATGTGAT 1320
Db 1052 LeuSerProAlaLeuGlyLeuProAspIleThrLysProPheGluLeuPheIleAsp 1071
Qy 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCCAAACCTAGGACCATGAGAGACCT 1380
Db 1072 GluAsnSerGlyPheAlaLysGlyValLeuValGlnLysLeuGlyProTrpLysArgPro 1091
Qy 1381 GTCGCCTACCTGTCAAGAAGCTCGATCCTGTAGCCAGTGTGTGGCCCATATGCTGAAG 1440
Db 1092 ValAlaTyrLeuSerLysLeuAspThrValAlaSerGlyTrpProProCysLeuArg 1111
Qy 1441 GCTATCGAGCTGTGGCCATCTGTCAGGACGCTGCAAAATTGACTTTGGGACAGAA 1500
Db 1112 MetValAlaAlaIleAlaIleLeuValLysAspAlaGlyLysLeuThrLeuGlyGlnPro 1131
Qy 1501 ATAACGTAAATAGCCCCCATGCTTGGAGAACATCGTTGCGAGCCCCCAGACCGATGG 1560
Db 1132 LeuThrIleLeuThrSerHisProValGluAlaLeuValArgGlnProProAsnLysTrp 1151
Qy 1561 ATGACCAACGCCCGCATGACCACTATCAAAAGCTGCTTCFC---ACAGAGAGGCTCAGC 1617
Db 1152 LeuSerAsnAlaArgMetThrHisTyrGlnAlaMetLeuLeuAspAlaGluArgValHis 1171
Qy 1618 TTCGCTCACCGCGCTCTCAACCTGCGCACTTTCTGCTGAGAGAGCTGATGAACCA 1677
Db 1172 PheGlyProThrValSerLeuAsnProAlaThrLeuLeuProLeuProSerGlyLysPro 1191
Qy 1678 -----GTGACTCATGATGCCATCACTATTGATTGAGGAGACTGGGGTCCGCAAGCAG 1731
Db 1192 ProArgLeuSerProAsp-----LeuAlaGluThrMetAlaGlnThrAsp 1206
Qy 1732 CTTCACAGACATACCGCTGACTGGAGAAAGTGTAACTCTGGTTTCACTGACGGAAGCAGCTAT 1791
Db 1207 LeuThrAspGlnProLeuProAspAlaAspLeuThrTyrThrAspGlySerSerPhe 1226
Qy 1792 GTGGTGAAGATAGAGATGGCTGGGGCGCGGTGTGGACGGGACCGCACCATCTGG 1851
Db 1227 IleArgAsnGlyGluArgLysAlaGlyAlaAlaValThrThrGluSerGluValIleTrp 1246
Qy 1852 GCCAGACGCTCGCGGAAGGAACCTTCAGCACAAGGCTGACTGCGCCCTCACGCAA 1911
Db 1247 AlaAlaSerLeuProProGlyThrSerAlaGlnArgAlaGluLeuIleAlaLeuThrGln 1266
Qy 1912 GCTTTGCGGCTGGCCGAAGGNAATCCATAAACATTTATACGGACAGCAGTATGCTTT 1971
Db 1267 AlaLeuLysMetAlaLysGlyLysLeuThrValTyrThrAspSerArgTyrAlaPhe 1286
Qy 1972 GCGACTGCACATGATGGGCGCATCTATAACAAAGGGGGTGTGCTTACTTCAGCAGGG 2031
Db 1287 AlaThrAlaHisValHisGlyGluIleTyrArgArgGlyLeuLeuThrSerGluGly 1306
Qy 2032 AGGGAATAAGAACAAAGAGGAATTTAAAGCCTATTAGACCGGTACATTTACCAAAA 2091
Db 1307 LysGluIleLysAsnLysAsnGluIleLeuAlaLeuLeuGluAlaLeuPheLeuProLys 1326
Qy 2092 AGGTAGCTATTATACATGCTCTGGACATCAGAAAGCTAAGATCTCATATCCAGAGGA 2151
Db 1327 ArgLeuSerIleIleHisCysProGlyHisGlnLysGlyAspSerProGlnAlaLysGly 1346
Qy 2152 AACCATGCTGACCGGGTTCACAGCAGGCAGCC-----CAGGGGTGTTAAC 2199
Db 1347 AsnArgLeuAlaAspAspThrAlaLysLysAlaAlaThrGluThrGlnSerSerLeuThr 1366
Qy 2200 CTTCCTGCT-----ATAATAGAAATGCCCAAGCCCCCAGAACCCAGACAGCTACACC 2253
Db 1367 IleLeuProThrGluLeuIleGluGlyProLys-----ArgProProTrpGluTyrAsp 1384

Qy 2254 CTAGAGAGCTGGCAAGAGATAAAAGATAGACAGAGTTCTCTGAGACTCCGGAAGGGACC 2313
Db 1385 AspSerAspLeuAspLeuValGlnLysLeuGluAlaHisTyrGluProLysArgGlyThr 1404
Qy 2314 TGTATACCTCAGATGGGAAGAAATCTGCGCCCAACAAAGGGTTAGAAATATGTCOA 2373
Db 1405 ---TrpGluTyrArgGlyLysThrIleMetProGluLysTyrAlaLysGluLeuIleSer 1423
Qy 2374 CAGATACATCGTCTAACCCACCTAGGAACATAACACCTGCAGCAGTGTGGTCCAGAACATCC 2433
Db 1424 HisLeuHisLysLeuThrHisLeuSerAlaArgLysMetLysThrLeuLeuGluArgGlu 1443
Qy 2434 -----CCTTATCATGTTCTGAGGGCTACCCAGAGGTGGCTGACTCGGTG 2475
Db 1444 GluThrGlyPheTyrLeuProAsnArgAspLeuHisLeuArgGlnValThrGluSer--- 1462
Qy 2476 GTCMAACATTTGTGTGCCCTGCGAGCTGGTTAATGCTAATCTTCCAGAAATGCTCCAGGG 2535
Db 1463 -----CysArgAlaCysAlaGlnIleAsnAlaGlyLysIleLysPheGlyLeuAsp 1479
Qy 2536 AAGAGACTAAGGGGAAGCCACCCAGGCGCTCAGTGGGAAGTGGACTTCAGTGAGTAAAG 2595
Db 1480 ValArgAlaArgGlyArgArgProGlyThrHisTrpGluValAspPheThrGluIleLys 1499
Qy 2596 CCGGCTAAATACGGAAACAAATACCTATTGTTGTTTGTAGACACTTTTCAGAGTGGGTA 2655
Db 1500 ProGlyMetTyrGlyTyrLysTyrLeuLeuValPheIleAspThrPheSerGlyTrpAla 1519
Qy 2656 GAGCTTATCTACTAAGAAGAGACTTCAACCGTGTGGCTAAAAAATACTGGAAGAA 2715
Db 1520 GluAlaTyrProAlaLysHisGluThrAlaLysValValAlaLysLysLeuGluGlu 1539
Qy 2716 ATTTTCCAGATTTGGAATACCTAAGTAATAGGTAGGTGAGACATGTCAGCTTTGTT 2775
Db 1540 IlePheProArgTyrGlyIleProGlnValLeuLysSerAspAsnGlyProAlaPheIle 1559
Qy 2776 GCCAGGTAAGTCAGGACTGGCCAAATATTGGGGATTTGAGGAACTGATTTGTCATGTCGA 2835
Db 1560 SerGlnValSerGlnSerValAlaThrLeuLeuGlyIleAsnTrpLysLeuHisCysAla 1579
Qy 2836 TACAGACCCCAAGCTCAGGACAGGTAGAGAGATGATGAACCATTAAGAGACCTT 2895
Db 1580 TyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgSerIleLysGluThrLeu 1599
Qy 2896 ACTAAATTGACCGCGAGACTGGCGTTAATGATTGGATAGCTCTCTCCCTGCTTTGCTT 2955
Db 1600 ThrLysLeuThrLeuGluThrGlySerLysAspTrpValLeuLeuLeuProLeuValLeu 1619
Qy 2956 TTTAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGTAATTTACTTACGGG 3015
Db 1620 TyrArgValArgAsnThrProGlyProHisGlyLeuThrProPheGluIleLeuTyrGly 1639
Qy 3016 GGACCCCGCCCATTTGTTAGAAATTTGCTTCTGTATAGTGTGACGTGCTGCTTTCCAG 3075
Db 1640 AlaProProMetAlaHisPhePheAspThrAspIleSerThrPheAlaThrSerPro 1659
Qy 3076 CCTTTGTTCTTAGGCTCAAGGCATTTGAGTGGGTGAGACACAGAGCGTGGAGGCACTC 3135
Db 1660 ThrMetGlnAlaHisLeuArgAlaLeuGlnLeuValGlnGluGluIleGlnArgProLeu 1679
Qy 3136 CGGAGGCGCTTACTCAGGAGGAGGAGACTTCGAG---ATCCCAACATCTTTCCAGTGGGA 3192
Db 1680 AlaAlaAlaTyrArgGluLysLeuGluThrProValValProHisProPheLysProGly 1699
Qy 3193 GATTCACTGCTTACGTTAGACGCCCGGTGAGGAAACCTCGAGACTCGGTGGAAGGGCCCT 3252
Db 1700 AspSerValTrpValArgArgHisGlnThrLysAsnLeuGluProArgTrpLysGlyPro 1719
Qy 3253 TATCTGCTACTTTTACCACCAACCGGCTGTGAAGTGAAGGAATCTCCACTGATC 3312
Db 1720 HisIleValLeuLeuThrThrProThrAlaLeuLysValAspGlyValAlaAlaTrpIle 1739

```
Qy 3313 CATGATCCACGTTAAACCGCGCCACCT----- 3342
Db ||||| ||||| ||||| ||||| ||||| |||||
1740 HisAlaSerHisValysAlaAlaGlyProThrThrAsnGlnAspLeuSerAspSerPro 1759
Qy 3343 -----CCCGATTTCGGGTGGAAAGCCGAAAGACTGAAAGACTCCCTTTAAAGCTTCGC 3393
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1760 SerSerAspAspProSerArgTrpIysValGlnArgThrGlnAsnProLeuLysIleArg 1779
Qy 3394 CTCATCGC 3402
Db ||||| |||||
1780 LeuSerArg 1782

RESULT 14
AAAY03163
ID AAY03163 standard; protein; 1079 AA.
AC AAY03163;
XX
DT 10-JUN-1999 (first entry)
XX
DE MuLV reverse transcriptase.
XX
KW MuLV reverse transcriptase; dscDNA production; cDNA library production.
XX
OS Murine leukemia virus.
XX
PN US8891637-A.
XX
PD 06-APR-1999.
XX
PF 15-SEP-1997; 97US-00929967.
XX
PR 15-OCT-1996; 96US-00126428.
XX
PA (GETH ) GENENTECH INC.
XX
PI Rupert SJ;
XX
DR WPI; 1999-253852/21.
XX
PT Producing double stranded cDNA molecules from mRNA transcripts and a
PT novel xenotropic murine leukemia virus reverse transcriptase - useful for
PT constructing cDNA libraries.
XX
PS Claim 32; Col 45-52; 70pp; English.
XX
CC This sequence represents a murine leukaemia virus (MuLV) reverse
CC transcriptase. The invention relates to a method of producing double
CC stranded cDNA (dscDNA) molecules from mRNA transcripts and a xenotropic
CC murine leukaemia virus reverse transcriptase (XM-MuLV). The method, the
CC host cells, reverse transcriptases and nucleic acids disclosed may be
CC used to produce high quality, normalized, full-length, directionally
CC cloned, cDNA libraries. Use of the method to construct cDNA libraries
CC ensures that the reverse transcriptase is able to extend the first strand of
CC the cDNA up to the ultimate 5' end of the mRNA, therefore ensuring that
CC all sequences are represented in the library, whereas previously, some
CC sequences would be lost due to incomplete transcription (e.g. as a result
CC of hairpin loop formation) or mRNA degradation (e.g. as a result of RNase
CC H activity). Also, carrying the process out in cells means that enzymes
CC and proteins present in the cell will repair any mistakes or nicks in the
CC product DNA. Modifying the mRNA for insertion into a vector by adding the
CC 5' cap provides a convenient way of priming it for cDNA synthesis. Also,
CC the method provides a fast and clean way of synthesizing cDNA clones
XX
SQ Sequence 1079 AA;

Alignment Scores:
Pred. No.: 7.6e-312 Length: 1079
Score: 3600.00 Matches: 679
Percent Similarity: 76.7% Conservative: 156
Best Local Similarity: 62.4% Mismatches: 222
Query Match: 58.2% Indels: 32
DB: 2 Gaps: 11
```

```
US-10-723-552-3_COPY_2307_5741 (1-3435) x AAY03163 (1-1079)
Qy 211 ACCCTCCAATAGATGAGAAATATCGACTATCTCTCCCTAGTAAAGCGTGATCAAAAT 270
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 ThrLeuAsnIleGluAspGluHisArgLeuHisGluThrSerLysGluProAspValSer 20
Qy 271 ATA---CAATCTCGTTGGAACAGTTTCCCGACAGCTGGCGAGAAACCGCAGGATCGGT 327
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 LeuGlySerThrTrpLeuSerAspPheProGlnAlaTrpAlaGluThrGlyGlyMetGly 40
Qy 328 TTGGCAAGCAAGTTCCCCACACAGTTATTCAACTGAAGGCCAGTGCACACACAGTGTCA 387
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 LeuAlaValArgGlnAlaProLeuIleProLeuLysAlaThrSerThrProValSer 60
Qy 388 GTCAGACAGTACCCCTTGAGTAAGAAGCTCAAGAGAAATTCGGCCGCATGTCCAAGA 447
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 IleLysGlnTyrProMetSerGlnAlaArgLeuGlyIleLysProHisIleGlnArg 80
Qy 448 TTAATCCACAGGGCATCTAGTTCCTGCTCAATCTCCCTGGAATACTCCCTCTCTACCG 507
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 LeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeuLeuPro 100
Qy 508 GTTAGAAAGCCTGGGACTAATGACTATCGACCACTACAGGACTTGAGAGAGGTCAATAA 567
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 ValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLys 120
Qy 568 CGGTGTCAGGATATACACCAACAGTCCCGAACCTTATACCTTGTGTGCTCTCCCA 627
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 ArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeuPro 140
Qy 628 CCCCAACGGAGCTGGTATACAGTATTGAGCTTAAAGGATGCCTTCTCTGCTCAGATTA 687
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
141 ProSerHisGlnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeu 160
Qy 688 CACCCCACTAGCCAAACCACTTTTGCCTTGAATGGAGAGATCCAGGTACGGGAAGAAC 747
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
161 HisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGluMetGlyIleSer 180
Qy 748 GGGCAGCTACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCGACCATCTTTGAC 807
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 GlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAsp 200
Qy 808 GAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCAAACACCTCAGGTGACCTCTCTC 867
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 GluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuIleLeuLeu 220
Qy 868 CAGTACGTGGATGACCTGCTTCTCGCGGGAGCCCAACAAACAGGACTGCTTAGAAGGCACG 927
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
221 GlnTyrValAspAspLeuLeuAlaAlaThrSerGluLeuAspCysGlnGlnGlyThr 240
Qy 928 AAGGACTACTGCTGGGAATTTCTGACCTAGGCTACAGAGCTCTGCTAAGAGCCCGCAG 987
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 ArgAlaLeuLeuGlnThrLeuGlyAsnLeuGlyTyrArgAlaSerAlaLysLysAlaGln 260
Qy 988 ATTTTCAGGAGAGAGGTAAACATACCTGGGGTACAGTTTGGCGGACGGSCAGCATGGCTG 1047
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
261 IleCysGlnLysGlnValLysTyrLeuGlyTyrLeuLeuLysGluGlyGlnArgTrpLeu 280
Qy 1048 ACGGAGGCACCGAAGAAAACCTGTAGTCCAGATACCGGCCCCCAACACACGCAACAAATG 1107
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
281 ThrGluAlaArgLysGluThrValMetGlyGlnProThrProLysThrProArgGlnLeu 300
Qy 1108 AGAGAGTTTTGGGGACAGCTGGATTTTGGAGACTGTGGATCCCGGGGTTCGCGACCTTA 1167
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 ArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGluMet 320
Qy 1168 GCAGCCCACTCTACCCGCTAAACCAAGAAAAGGGGAATTCCTCGGGCTCTTGAGCAC 1227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
321 AlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAspGln 340
Qy 1228 CAGAAGGCATTGTGATGCTATCAAAAAGGCCCTCTGCTGAGCGCACCTGCTCTGGCCCTCCCT 1287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 341 GlnLysAlaTyrGlnGluIleLysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeuPro 360
Qy 1288 GACGTAACTAAACCTTTACCTTTATGTGGATGAGCGTAAGGAGTAGCCCGGGAGTT 1347
Db 361 AspLeuThrLysProPheGluLeuPheValAspGluLysGlnGlyTyrAlaLysGlyVal 380
Qy 1348 TTAACCCAAACCCTAGGACCAATGAGAAAGACTGTGTGGCTACCTGTCTCAAGAAGCTCAT 1407
Db 381 LeuThrGlnLysLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAsp 400
Qy 1408 CCTGTAGCCAGTGGTGGCCCATATGCTCGAAGGCTATCGCAGCTGTGGCCATCTGGTC 1467
Db 401 ProValAlaAlaGlyTrpProCysLeuArgMetValAlaAlaIleAlaValLeuThr 420
Qy 1468 AAGGACGCTGCAAAATGACTTTGGGACAGAAATATACTGTAATAGCCCCCATGCATTTG 1527
Db 421 LysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAlaVal 440
Qy 1528 GAGAACATCGTTCCGGCAGCCCGACGACCGATGATGATGATGATGATGATGATGATGAT 1587
Db 441 GluAlaLeuValLysGlnProProAspArgTrpLeuSerAsnAlaArgMetThrHisTyr 460
Qy 1588 CAAAGCCTGCTTCTC---ACAGAGAGGTCACGTTTCGCTCCACAGCGCTCTCAACCT 1644
Db 461 GlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProValValAlaLeuAsnPro 480
Qy 1645 GCCACTCTTCTGCTGAAGACTGATGAACAGTGAACAGTGAACAGTGAACAGTGAAC 1704
Db 481 AlaThrLeuLeuPro---LeuProGluGluGlyLeuGlnHisAsnCysLeuAspIleLeu 499
Qy 1705 ATTGAGAGACTGGGTCGCGACGACCTTACAGACATACCGCTGACTCGGAGAGTGCTA 1764
Db 500 AlaGluAlaHisGlyThrArgProAspLeuThrAspGlnProLeuProAspAlaAspHis 519
Qy 1765 ACCTGGTTCACTGACGGAAGCAGTATGTGTGTAAGTAAGGATGCTGGCGCGCG 1824
Db 520 ThrTrpTyrThrAspGlySerSerLeuLeuGlnGlyGlnArgLysAlaGlyAlaAla 539
Qy 1825 GTGTGACGCGGACCCGACCATCTGGGCGACGACCTGCGGGAAGAACTTCAGCACAA 1884
Db 540 ValThrThrGluThrGluValIleTrpAlaLysAlaLeuProAlaGlyThrSerAlaGln 559
Qy 1885 AAGCTGAGCTATGCGCCTCACGCAAGCTTTGGCGTGGCGGAGGGAATCCATAAAC 1944
Db 560 ArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysLysLeuAsn 579
Qy 1945 ATTTATACGACGACGAGTATGCTTTCGACTGCACAGTACATGCGGCGCCATCTATAA 2004
Db 580 ValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyGluIleTyrArg 599
Qy 2005 CAAAGGGGTGTGCTTACCTCAGCAGGAGGGAATAAAGAACAAAGGAAATCTTAAGC 2064
Db 600 ArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAspGluIleLeuAla 619
Qy 2065 CTATTAGACCGGTACATTTACCAAAAGGCTAGCTATTATACACTGCTCTGACATCAG 2124
Db 620 LeuLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyHisGln 639
Qy 2125 AAAGCTAAAGATCTCATATCCAGAGGAACCAAGTGGCTGACCGGGTTGCCAGCAGGCA 2184
Db 640 LysGlyHisSerAlaGluAlaArgGlyAsnArgMetAlaAspGlnAlaAlaArgLysAla 659
Qy 2185 GCCCAGGGTGTAAACCTTCTGCTCTATATAAGAAATGCCAAAGCC----- 2229
Db 660 Ala-----IleThrGluThrProAspThrSerThrLeuLeuIle 672
Qy 2230 -----CCAGAACCCAGACAGTACACCTTAGAAGACTGGCAAGAG 2271
Db 673 GluAsnSerSerProTyrThrSerGluHisPheHisTyrThrValThrAspIleLysAsp 692
Qy 2272 ATAAAAGATAGACACG---TTCTCTGACACTCCGGAAGGACCTGCTATACCTCAGAT 2328
Db 693 LeuThrLysLeuGlyAlaIleTyrAspLysThrLysLysTyrTrpValTyr-----Gln 710

Qy 2329 GGGAGGAAATCTCTGCCCAACAAAGAGGGTTAGAAATATGTCCAACAGATATCGTCTA 2388
Db 711 GlyLysProValMetProAspGlnPheThrPheGluLeuLeuAspPheLeuHisGlnLeu 730
Qy 2389 ACCCACCTAGGAATAAACACCTGCAGCAGTGTGTGCAGAACCA-----TCCCTTATCAT 2442
Db 731 ThrHisLeuSerPheSerLysMetLysAlaLeuLeuGluArgSerHisSerProTyrTyr 750
Qy 2443 GTTCTGAGGCTACCAGGAGTGGCTGACTCGTGTGTCAAACTGTGTGCTCCCTGCCAGCTG 2502
Db 751 MetLeuAsnArgAspArgThrLeuLysAsnIleThrGluThrCysLysAlaCysAlaGln 770
Qy 2503 GTTAATGCTAATCTCTCCAGAAATGCTCCAGGGAAGAGACTAAAGGGAAGCCACCCAGGC 2562
Db 771 ValAsnAlaSerLysSerAlaValLysGlnGlyThrArgValaArgGlyHisArgProGly 790
Qy 2563 GCTCAGCTGGGAAGTGGACTTCACTGAGGTAAAGCCGCTAAATACGGAACAATAATACCTA 2622
Db 791 ThrHisTrpGluIleAspPheThrGluIleLysProGlyLeuTyrGlyTyrLysTyrLeu 810
Qy 2623 TTGTTTTTTGTAGACACCTTTTTCAGGATGGTAGAGCTTATCTTACTTAAGAAAGAGACT 2682
Db 811 LeuValPheIleAspThrPheSerGlyTrpIleGluAlaPheProThrLysLysGluThr 830
Qy 2683 TCACCCGTGTGGCTAAAGAAATACTCGAAGAAATTTTCCAGATTTTGAATACCTTAAG 2742
Db 831 AlaLysValValThrLysLysLeuLeuGluGluIlePheProArgPheGlyMetProGln 850
Qy 2743 GTAATAGGTCAGACAAATGTCAGCTTTTGTGTCGCCAGGTAACTAGGACTCGCCCAAG 2802
Db 851 ValLeuGlyThrAspAsnGlyProAlaPheValSerLysValSerGlnThrValAlaAsp 870
Qy 2803 ATATTGGGATTCATTTGGAACCTTACTTAAATTTGACCGCGAGACTGGCGTT 2922
Db 891 GluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuAlaThrGlySer 910
Qy 2923 AATGATTGGATAGCTCTCTGCGCTTTGTGCTTTTATAGGTTAGAACACCCCTGGACAG 2982
Db 911 ArgAspTrpValLeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThrProGlyPro 930
Qy 2983 TTTGGGCTGACCCCTATGAAATTTACTCTACGGGGGACCCCTTGGTAGAAATTCCT 3042
Db 931 HisGlyLeuThrProTyrGluIleLeuTyrGlyAlaProProProLeuValAsnPhePro 950
Qy 3043 TCTGTACATAGTCTGACGTGCTTTCAGCCTTTTCTTAGCTTCTAGGCTCAAGGCACTT 3102
Db 951 AspProAspMetThrArgValThrAsnSerProSerLeuGlnAlaHisLeuGlnAlaLeu 970
Qy 3103 GAGTGGGTGAGACACAGCGTGGAGCAACTCGGAGGCGCTACTCTCAGGAGGAGGAGAC 3162
Db 971 TyrLeuValGlnHisGluValTrpArgProLeuAlaAlaAlaTyrGlnGluLeuAsp 990
Qy 3163 TTGCAG---ATCCACATCGTTTCCAAGTGGGAGATTCAGTCTACGTGTAGCCACCGCT 3219
Db 991 ArgProValValProHisProTyrArgValGlyAspThrValTrpValaArgHisGln 1010
Qy 3220 GCAGGAACCTCCAGACTCGGTGGAAGGCGCTTATCTCGTACTTTTGGACACCAACG 3279
Db 1011 ThrLysAsnLeuGluProArgTrpLysGlyProTyrThrValLeuLeuThrThrProThr 1030
Qy 3280 GCTGTGAAGTGAAGAAATCTCCACTGGATCCATGCATCCACGTTAAACCGGCGCCA 3339
Db 1031 AlaLeuLysValAspGlyIleAlaAlaTyrIleHisAlaAlaHisValLysAlaAlaAsp 1050
Qy 3340 CCT-----CCCCATTCG-----GGTGAAGAGCCGAAAGACTGAAATCCCTT 3384
Db 1051 ProGlyGlyProSerArgLeuThrTrpArgValGlnArgSerGlnAsnProLeu 1070


```

QY 3220 GCAGGAACCTCGAGACTCGGTGGAGGGCCCTTATCTCGTACTTTTGACCAGACCAACG 3279
Db 1129 PheGlnGlnLeuGlyProArgTTPAspGlyProTyrThrValValLeuSerThrProThr 1148
QY 3280 GCTGTGAAGTCGAAGGAATCTCCACCTGGATCCATGCCACGTTAAACCGCGCCA 3339
Db 1149 AlaValLysValAlaGlyLysThrProTrpIleHisSerArgLeuLysLysAlaPro 1168
QY 3340 CCTCCCGATTCCGGGTGGAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCCTCCAT 3399
Db 1169 AspAsnGlnGluGluTyrThrValSerProThrSerAspProLeuArgValLysLeuThr 1188
QY 3400 CGC 3402
Db 1189 Arg 1189

```

Search completed: February 14, 2006, 15:54:26
Job time : 670.958 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 15:55:36 ; Search time 17.7535 Seconds
(without alignments)
3199.272 Million cell updates/sec

Title: US-10-723-552-3_COPY_2307_5741

Perfect score: 6183

Sequence: 1 ATGGGTGCCAGGCAACA.....CTGTCAATAACCTCTCAGAC 3435

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p_model -DEV=xlp
-Q=/abs/ABSSWEB/spool/US10723552/runat 14022006 135146 13079/app query.fasta.1
-DB=Issued_Patents_AA -QMT=fascan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs802p -USER=US10723552 @CGN 1.1 161 @runat 14022006 125146 13079
-NCPU=6 -ICPU=3 -NO_WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/H COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/PCUTS COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5851.5	94.6	1194	2	US-09-171-553B-5
2	4143.5	67.0	1203	2	US-09-075-272-4
3	3797.5	61.4	1737	2	US-09-309-572-13
4	3797.5	61.4	1737	2	US-09-718-096-13
5	3600	58.2	1079	1	US-08-929-967-8
6	2500.5	40.4	665	1	US-08-929-967-7
7	1501.5	24.3	379	2	US-09-603-185-6
8	1415.5	22.9	768	2	US-08-979-847B-89
9	1390.5	22.5	683	2	US-08-979-847B-198
10	1390.5	22.5	683	2	US-08-979-847B-208
11	1387.5	22.4	683	2	US-08-979-847B-208
12	1387.5	22.4	683	2	US-08-979-847B-210

13	1331.5	21.5	654	2	US-08-979-847B-91
14	1236	20.0	768	2	US-09-120-653D-5
15	629.5	10.2	917	2	US-08-259-451-11
16	579	9.4	995	4	PCT-US95-04910-14
17	578.5	9.4	1802	2	US-03-322-478-18
18	578.5	9.4	1802	2	US-09-586-106D-18
19	578.5	9.4	1802	2	US-10-799-870-18
20	528	8.5	1150	2	US-09-238-303-9
21	528	8.5	1150	2	US-09-946-239-9
22	512.5	8.3	1055	2	US-08-659-251-5
23	512.5	8.3	1055	2	US-09-256-490-5
24	512.5	8.3	1055	4	PCT-US96-11445-5
25	512	8.3	1122	2	US-09-248-796A-16374
26	508	8.2	1031	2	US-08-811-682-15
27	506.5	8.2	1014	2	US-09-319-588C-6
28	493.5	8.0	858	2	US-09-960-428-7
29	491	7.9	1018	2	US-09-206-551-46
30	487.5	7.9	3080	6	5223423-4
31	473.5	7.7	1016	2	US-09-625-972-24
32	470	7.6	1003	2	US-10-290-579A-193
33	460	7.4	1002	2	US-10-290-579A-189
34	460	7.4	1003	1	US-07-743-357-9
35	460	7.4	1005	2	US-10-290-579A-191
36	458.5	7.4	1003	2	US-10-290-579A-192
37	455.5	7.4	1003	2	US-10-290-579A-183
38	452.5	7.3	913	1	US-07-743-357-6
39	450	7.3	1000	2	US-10-290-579A-186
40	449	7.3	999	2	US-10-290-579A-190
41	446	7.2	1003	1	US-07-743-357-8
42	446	7.2	1004	1	US-07-743-357-7
43	443	7.2	1003	2	US-09-393-795-9
44	443	7.2	1003	2	US-09-393-795-11
45	440	7.1	1003	1	US-07-743-357-10

ALIGNMENTS

US-09-171-553B-5
Sequence 5, Appli
Patent No. 6756227
GENERAL INFORMATION:
APPLICANT: GALBRAITH, DANIEL N.
APPLICANT: HAWORTH, CHRISTINE
APPLICANT: LEES, GILLIAN M.
APPLICANT: SMITH, KENNETH T.
TITLE OF INVENTION: PORCINE RETROVIRUS
FILE REFERENCE: CFV-5.01
CURRENT APPLICATION NUMBER: US709/171.553B
CURRENT FILING DATE: 1999-02-08
PRIOR APPLICATION NUMBER: PCT/GB97/01087
PRIOR FILING DATE: 1997-04-18
PRIOR APPLICATION NUMBER: GB 9702668.6
PRIOR FILING DATE: 1997-02-10
PRIOR APPLICATION NUMBER: GB 9608164.1
PRIOR FILING DATE: 1996-04-19
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1194
TYPE: PRT
ORGANISM: Porcine retrovirus
US-09-171-553B-5

Alignment Scores:
Pred. No.: 0
Score: 5851.50
Percent Similarity: 98.4%
Best Local Similarity: 96.9%
Query Match: 94.6%
DB: 2
Length: 1194
Matches: 1107
Conservative: 17
Mismatch: 17
Indels: 1
Gaps: 1

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-09-171-553B-5 (1-1194)

QY 1 ATGGGTGCCACAGGCAACACAGTATCCATGGACTACCCGAGAACAGTTGACTTGGGA 60
Db 50 MetGlyAlaThrGlyGlnArgGlnProTrpThrThrArgThrValAspLeuGly 69
QY 61 GTGGAGCGGGTAACCCACTCGCTTTCTGTGTATACCTGAGTGCACCCAGCACCCCTCTTAGGT 120
Db 70 ValGlyArgValThrHisSerPheLeuValIleProGluCysProValProLeuLeuGly 89
QY 121 AGAGACTTATTGACCAAGATGGGAGCACAAATTTCTTTTGAAACAGGGAACCAAGAGTG 180
Db 90 ArgAspLeuLeuThrLysMetGlyAlaGlnIleSerPheGluGlnGlyArgProGluVal 109
QY 181 TCTGCAATACAAACCTATCACTGTGTGACCTCCCAATTAGATGAGATATCGACTA 240
Db 110 SerValAsnAsnLysProIleThrValLeuThrLeuGlnLeuAspAspGluTyArgLeu 129
QY 241 TACTCTCCCTAGTAAGCCTGATCAAAATATACAAATTTCTGTTGGAAACAGTTTCCCCAA 300
Db 130 TyrSerProGlnValLysProAspGlnAspIleGlnSerTrpLeuGluGlnPheProGln 149
QY 301 GCCTGGCAGAAACCGCAGGATGGTTTGCAAGCAAGTTCCCCCAAAAGTTATTCAA 360
Db 150 AlaTrpAlaGluThrAlaGlyMetGlyLeuAlaLysGlnValProProGlnValIleGln 169
QY 361 CTGAAGCCAGTCCACACAGTGTCACTGAGTCAGTACAGTACCCCTTGAGTAAGAAGCTCAA 420
Db 170 LeuLysAlaSerAlaThrProValSerValArgGlnTyrProLeuSerArgGluAlaArg 189
QY 421 GAAGGAATTCGGCGCATGTCCAAAGATTAAATCAACAGGCATCCTAGTTCTCTGTCCAA 480
Db 190 GluGlyIleTrpProHisValGlnArgLeuIleGlnGlnGlyIleLeuValProValGln 209
QY 481 TCTCCCTGGANATCTCCCTGCTACCGTTAGAAAGCCTGGGACTAATGACTATCGACCA 540
Db 210 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyArgPro 229
QY 541 GTACAGGACTTGAAGAGGTCAATAACGGTGCAGGATATACCCCAACAGTCCCGAAC 600
Db 230 ValGlnAspLeuArgGluValAsnLysArgValGlnAspIleHisProThrValProAsn 249
QY 601 CTTATACCTCTTGTGTGCTCTCCACCCCAAGGAGCTGTGTATACAGTATTGGACTTA 660
Db 250 ProTyrAsnLeuLeuSerAlaLeuProProGluArgAsnTrpTyrThrValLeuAspLeu 269
QY 661 AAGGATGCTTCTTGTGCTGAGATTACACCCCACTAGCCAAACCACTTTTTCCTTGGAA 720
Db 270 LysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPheGlu 289
QY 721 TGGAGAGATCCAGGTACGGGAGAACCGGGCAGCTCACCTGGAACCGGCTCCCAAGG 780
Db 290 TrpArgAspProGlyThrGlyArgThrGlyGlnLeuThrTrpThrArgLeuProGlnGly 309
QY 781 TTCGAAGACTCCCGACCATCTTTGACGAAGCCCTTACACAGAGACTGCGCCAACTTCAGG 840
Db 310 PheLysAsnSerProThrIlePheAspGluAlaLeuHisArgAspLeuAlaAsnPheArg 329
QY 841 ATCCAACACCTCAGGTACCCCTCTCCAGTACGTGGATGACCTGCTTCTGCGGGAGCC 900
Db 330 IleGlnHisProGlnValThrLeuLeuGlnTyrValAspAspLeuLeuLeuAlaGlyAla 349
QY 901 ACCAACAGGACTGCTTAGAGGCACGAAGGCACTACTGCTGGATTTCTGACCTAGGC 960
Db 350 ThrLysGlnAspCysLeuGluGlyThrLysAlaLeuLeuLeuGluLeuLeuSerAspLeuGly 369
QY 961 TACAGAGCTCTGCTAAGAAGCCAGATTTGACGAGAGAGGTAAACATACTTGGGGTAC 1020
Db 370 TyrArgAlaSerAlaLysLysAlaGlnIleCysArgArgGluValThrTyrLeuGlyTyr 389
QY 1021 AGTTTGGGGACGGCAGCGATGGCTGACGAGGCGACGAAGAAACCTGTAGTCCAGATA 1080
Db 390 SerLeuArgGlyGlyGlnArgTrpLeuThrGluAlaArgLysLysThrValValGlnIle 409

QY 1081 CCGGCCCAACCAACAGCCAAACAAATGAGAGAGTTTTTTGGGGAAGTGGATTTTTCAGA 1140
Db 410 ProAlaProThrThrAlaLysGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 429
QY 1141 CTGTGGATCCCGGGTTTGGACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAAA 1200
Db 430 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrLysGluLys 449
QY 1201 GGGGAATTCCTCGGGCTCTGAGCACCAAGGCAATTTGATGCTATCAAAAAGGCCCTG 1260
Db 450 GlyGlyPheSerTrpAlaProGluHisGlnLysAlaPheAspAlaIleLysLysAlaLeu 469
QY 1261 CTGAGCCACACTGCTCGGCCCTCCCTCAGCTAACTAAACCTTTTACCTTTATGTGGAT 1320
Db 470 LeuSerAlaProAlaLeuAlaLeuProAspValThrLysProPheThrLeuTyrValAsp 489
QY 1321 GAGCGTAAGGAGTAGCCCGGGAGTTTAAACCAACCTTAGGACCATGGAGAACCT 1380
Db 490 GluArgLysGlyValAlaArgGlyValLeuThrGlnThrLeuGlyProTrpArgArgPro 509
QY 1381 GTCCCTACCTGTCAAGAAGCTCGATCCTGTAGCCAGTGGTTGGCCCATATGCTGAAG 1440
Db 510 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProValCysLeuLys 529
QY 1441 GCTATCCAGCTGTGGCCATCTGTCAGGACGCTGACAAATGACTTTGGGACAGAAT 1500
Db 530 AlaIleAlaAlaValAlaIleLeuValLysAspAlaAspLysLeuThrLeuGlyGlnAsn 549
QY 1501 ATAACTGTAATAGCCCCCATGTCATGAGAAACATCTTGGCAGCCCCAGACCGATGG 1560
Db 550 IleThrValIleAlaProHisAlaLeuGluAsnIleValArgGlnProProAspArgTrp 569
QY 1561 ATGACCAACGCCCGCATGCCACTATCAAGCTGCTCTTCTCACAGAGAGGCTCACCTTC 1620
Db 570 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuThrGluArgValThrPhe 589
QY 1621 GCTCCACAGCCGCTCTCAACCTCCCACTCTTCTGCTGAAGAGACTGATGAACCACTG 1680
Db 590 AlaProProAlaAlaLeuAsnProAlaThrLeuLeuProGluGluThrAspGluProVal 609
QY 1681 ACTCATGATTGCCATCAACTATTGATTGAGAGACTGGGGTCCGACAGGACCTTACAGAC 1740
Db 610 ThrHisAspCysHisGlnLeuLeuIleGluGluThrGlyValArgLysAspLeuThrAsp 629
QY 1741 ATACCGCTGACTGAGAGTCTTAACCTGGTTCACTGACGGAAGCAGCTAGTGTGTGAA 1800
Db 630 IleProLeuThrGlyGluValLeuThrTrpPheThrAspGlySerSerTyrValValGlu 649
QY 1801 GGTAAAGAGATGGCTGGGGCGCGGTGTGTGACCGGACCCCGCAGCATCTGGGCCAGCAGC 1860
Db 650 GlyLysArgMetAlaGlyAlaAlaValValAspGlyThrArgThrIleTrpAlaSerSer 669
QY 1861 CTGCCGGAAGAACTTCAGCACAAAAGCTGAGCTCATGGCCCTCAGCCAAAGCTTTGGCG 1920
Db 670 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuMetAlaLeuThrGlnAlaLeuArg 689
QY 1921 CTGCCCGAGGGAATCCATAAACATTATACGACACAGGATGCTTTCGCACTGCA 1980
Db 690 LeuAlaGluGlyLysSerIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 709
QY 1981 CAGCTACATGGGGCATCTATAACAAAGGGGGTGTCTTACTCAGCAGGAGGAGGAATA 2040
Db 710 HisValHisGlyAlaIleTyrLysGlnArgGlyLeuLeuThrSerAlaGlyArgGluIle 729
QY 2041 AAGAACAAAGAGGAAATTTCTAAGCCTATTGAAGCCGTACATTTTCAAAAAGCGTAGCT 2100
Db 730 LysAsnLysGluGluIleLeuSerLeuLeuGluAlaLeuHisLeuProLysArgLeuAla 749
QY 2101 ATTATACACTGCTCTGGACATCAGAAAGCTTAAGATCTCATATCCAGGGAACCCAGATG 2160
Db 750 IleIleHisCysProGlyHisGlnLysAlaLysAspLeuIleSerArgGlyAsnGlnMet 769
QY 2161 GCTGACCGGGTTGCCAAAGCAGGCGCGGTGTAACTTCTGCTCTATTAATAGAAATG 2220

```
Db 770 AlaAspArgValAlaLysGlnAlaGlnAlaValAsnLeuLeuProIleIleGluThr 789
Qy 2221 CCCAAAGCCCGAAGAACCCAGACGACGATACACCTTAGAAGACTGGCAGAGATATAAAAG 2280
Db 790 ProLysAlaProGluProArgArgGlnThrLeuGluAspTrpGlnGluIleLysLys 809
Qy 2281 ATAGACCACTTCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGAAATC 2340
Db 810 IleAspGlnPheSerGluThrProGluGlyThrCysTyrThrSerTyrGlyLysGluIle 829
Qy 2341 CTGCCCCACCAAGAGGTTAGATATGTCCACAGATACATCTCTAAACCCACTAGGA 2400
Db 830 LeuProHisLysGluGlyLeuGluTyrValGlnGlnIleHisArgLeuThrHisLeuGly 849
Qy 2401 ACTAAACACCTGCAGCAGTGTGTCAGAACATCCCTTATCATGTTCTGAGGCTACCAAGGA 2460
Db 850 ThrLysHisLeuGlnGlnLeuValArgThrSerProTyrHisValLeuArgLeuProGly 869
Qy 2461 GTGCTCACTCGGTGGTCAACATTTGTGTGCCCTGCCAGCTGGTTAATGCTTAATCTCTCC 2520
Db 870 ValAlaAspSerValValLysHisCysValProCysGlnLeuValAsnAlaAsnProSer 889
Qy 2521 AGAATGCTCCAGGAGAGACTAAGGGAGACCCAGCGCTCACTGGGAAGTGCAC 2580
Db 890 ArgIleProProGlyLysArgLeuArgGlySerHisProGlyAlaHisTrpGluValAsp 909
Qy 2581 TTCACTCAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTTGTGTAGACACC 2640
Db 910 PheThrGluValLysProAlaLysTyrGlyAsnLysTyrLeuLeuValPheValAspThr 929
Qy 2641 TTTTCAGAGTGGTAGGCTTATCTCTAAGAAAGAGAGCTTCAACCGTGGTGGCTAAA 2700
Db 930 PheSerGlyTrpValGluAlaTyrProThrLysLysGluThrSerThrValValAlaLys 949
Qy 2701 AAAATACGGAGAAATTTTCCAGATTGGGAATACCTAAGTAAATAGGTCAGACAT 2760
Db 950 LysIleLeuGluGluIlePheProArgPheGlyIleProLysValIleGlySerAspAsn 969
Qy 2761 GGTCCAGCTTTGTTGCCAGGTAAAGTCAGGAGCTGCCAAGATATTGGGGATTGATTGG 2820
Db 970 GlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaLysIleLeuGlyIleAspTrp 989
Qy 2821 AAATGCAATTTGTGCATACAGACCCCAAGCTCAGACAGGTAGAGAGATGAATAGAAC 2880
Db 990 LysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThr 1009
Qy 2881 ATTAAGACAGCTTACTAATTCACCGCGAGACTGGCGTTAATGATTGATGCTCTC 2940
Db 1010 IleLysGluThrLeuThrLysLeuThrThrGluThrGlyIleAsnAspTrpMetAlaLeu 1029
Qy 2941 CTGCCCCCTTGTGCTTTTGGGTTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTAT 3000
Db 1030 LeuProPheValLeuPheArgValArgAsnThrProGlyGlnPheGlyLeuThrProTyr 1049
Qy 3001 GAATTAATCTACGGGGGACCCCGCCCAATTTGTTAGAAATTCCTCTGTACATAGTGTGAC 3060
Db 1050 GluLeuLeuTyrGlyGlyProProProLeuAlaGluIleAlaPheAlaHisSerAlaAsp 1069
Qy 3061 GTGCTGTTTCCAGCCCTTTGTTCTTAGGCTCAGGCACCTTGAAGTGGGTGAGAACGA 3120
Db 1070 ValLeuLeuSerGlnProLeuPheSerArgLeuLysAlaLeuGluTrpValArgGlnArg 1089
Qy 3121 GCGTGGAGGCACTCCGGGAGGCTTACTCAGGAGGAGACTTGCAGATCCACATCGT 3180
Db 1090 AlaTrpLysGlnLeuArgGluAlaTyrSer--GlyGlyAspLeuGlnValProHisArg 1108
Qy 3181 TTCAAGTGGAGATTCACTAGTTAGACGCGCACCGTCCAGAAACCTCGAGACTCGG 3240
Db 1109 PheGlnValGlyAspSerValTyrValArgArgHisArgAlaGlyAsnLeuGluThrArg 1128
Qy 3241 TGAAGGCGCTTATCTCGTACTTTTGAACACACCAACCGCTGTGAAGTGAAGGAATC 3300
```

```
Db 1129 TrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysValGluGlyIle 1148
Qy 3301 TCCACCTGGATCCATGATCCCGTTAAACCGCGCACCTCCCGATTGGGGTGGAAA 3360
Db 1149 ProThrTrpIleHisLysHisValLysProAlaProProAspSerGlyTrpArg 1168
Qy 3361 GCCGAAAGACTGAAATCCCTTAAGCTTGGCTCCATCGCGTTCCTTACTCTGTC 3420
Db 1169 AlaGluLysThrGluAsnProLeuLysLeuArgLeuHisArgLeuValProTyrSerAsn 1188
Qy 3421 AATAAC 3426
Db 1189 AsnAsn 1190

RESULT 2
US-09-075-272-4
; Sequence 4, Application US/09075272
; Patent No. 6136598
; GENERAL INFORMATION:
; APPLICANT: MILLER, A. DUSTY
; APPLICANT: WOLGAMOT, GREG
; APPLICANT: BONHAM, LYNN
; TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL
; TITLE OF INVENTION: PACKAGING CELL LINES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,272
; FILING DATE: 08-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,140
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: 14538A-003710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-075-272-4

Alignment Scores:
Pred. No.: 0 Length: 1203
Score: 4143.50 Matches: 786
Percent Similarity: 79.6% Conservative: 125
Best Local Similarity: 68.6% Mismatches: 213
Query Match: 67.0% Indels: 21
DB: 7

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-09-075-272-4 (1-1203)

Qy 1 ATGGGTGCCACAGGGCAACAGATATCCATGACATACCCGACAGACAGTTGACTTGGGA 60
Db 50 ValGlyAlaThrGlySerLysLeuTyrProTrpThrThrLysArgAlaLeuGlnIleAsp 69
```

QY 61 GTGGAGCGGTAAACCCACTCGTTCTTGCTCATCTAGTGGCCAGCACCCCTCTTAGGT 120
Db |||||
QY 70 LysAsnMetValThrHisSerPheLeuValIleProGluCysProAlaProLeuLeuGly 89
Db |||||
QY 121 AGAGACTTTATTGACCAAGATGGGACGACAAATTTCTTTTGAACAGGGGAAACCCAGAGTG 180
Db |||||
QY 90 ArgAspLeuLeuThrLysLeuLysAlaGlnValGlnPheThrSerGluGlyProGlnVal 109
Db |||||
QY 181 TCTGCAATAACAAACCTATCACTGTGTGTGACCCCTCCAAATTAGATGACGAATATCGACTA 240
Db |||||
QY 110 SerTrpGlyLysAlaProLeuAlaCysLeuValLeuSerThrGluGluGluTyrArgLeu 129
QY |||||
QY 241 TACTCTCCCTAGTAAGCCTGATCAAAATATACAAATTCGTGTGGACAGCTTCCCCAA 300
Db |||||
QY 130 HisGluGluGlnProLysGlyAlaAlaProLeuAsp----TrpValThrAlaPheProAsn 148
QY GCCTGGGAGAGAAACCGCAGGGATGGTTTGGCAAGCAAGTATCCCCCAACAAGTTATTTCAA 360
Db |||||
QY 149 ValTrpAlaGluGlnAlaGlyMetGlyLeuAlaLysGlnValProProValValValGlu 168
QY 361 CTGAAGCGCAGTGCACACAGTGTCACTCAGACAGTACCCCTTGAGTAAAGAGCTCAA 420
Db |||||
QY 169 LeuLysAlaAspAlaThrProIleSerValArgGlnTyrProMetSerLysGluAlaLys 188
QY |||||
QY 421 GAAGGAATTCGGCCGATGTCACAAAGATTATCCAAACAGGCGATCCTAGTTCTCTGCCAA 480
Db |||||
QY 189 GluGlyIleArgProHisIleArgLeuLeuLeuAspGlnGlyIleLeuValAlaCysGln 208
QY 481 TCTCCTCGGAATACTCCCTGCTACCGGTTAGAAAGCCTGGGACTTAATGACTATCGACCA 540
Db |||||
QY 209 SerProTrpAsnThrProLeuLeuProValArgLysProGlyThrAsnAspTyrArgPro 228
QY 541 GTACAGGACTTGAGAGAGGTCAATAAACGGGTGCAGGATATACACCCCAACAGTCCCGAAC 600
Db |||||
QY 229 ValGlnAspLeuArgGluValAsnLysArgValLeuAspIleHisProThrValProAsn 248
QY 601 CCTTATACCTCTGTGTGCTCTCCACCCCAACGGAGCTGGTATACAGTATTGGACTTA 660
Db |||||
QY 249 ProTyrAsnLeuLeuSerSerLeuProGluArgThrTrpTyrThrValLeuAspLeu 268
QY 661 AAGGATGCTCTTCTTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAA 720
Db |||||
QY 269 LysAspAlaPhePheCysLeuArgLeuHisProLysSerGlnLeuLeuPheAlaPheGlu 288
QY 721 TGGAGAGATCAGGTACGGGAGAACCGGGCGAGCTACCTGACCCGAGCTGCCCAAGGG 780
Db |||||
QY 289 TrpArgAspProGluGlyGlyGlnThrGlyGlnLeuThrTrpTyrArgLeuProGlnGly 308
QY 781 TTCAGAACTCCCGACCATCTTTCACCAAGCCCTACACAGAGACTCGCCCAACTTCAGG 840
Db |||||
QY 309 PheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaProPheArg 328
QY 841 ATCCAAACACCTCAGGTGACCCCTCTCCAGTACGTGGATGACCTGCTTGCGCGGAGCC 900
Db |||||
QY 329 AlaGlnAsnProGlnLeuThrLeuLeuGlnTyrValAspAspLeuLeuIleAlaAla 348
QY 901 ACCAAACAGGACTGCTTAGAGGCGCAAGGCACTACTCTGCGGAATTTGCTGACCTAGGC 960
Db |||||
QY 349 SerLysGluLeuCysGlnGlnGlyThrGluArgLeuLeuThrGluGlyAsnLeuGly 368
QY 961 TACAGAGCTCTGCTAAGAGGCCAGATTTGACGAGAGAGGTAAACATACTTGGGGTAC 1020
Db |||||
QY 369 TyrArgValSerAlaLysLysAlaGlnIleCysGlnThrGluValIleTyrLeuGlyTyr 388
QY 1021 AGTTTGGGGGAGCGGACGATGGCTGACGGAGGCGAGGAAACCTGTAGTCAGATA 1080
Db |||||
QY 389 ThrLeuArgGlyGlyLysArgTrpLeuThrGluAlaArgLysLysThrValMetMetIle 408
QY 1081 CCGGCCCCAACACAGCCCAACAAATAGAGAGATTTTGGGACAGCTGGATTTTGCAGA 1140
Db |||||
QY 409 ProProProThrThrProArgGlnValArgGluPheLeuGlyThrAlaGlyPheCysArg 428
QY 1141 CTGTGATCCCGGGGTTTGGACCTTAGCAGCCCTACTCTACCCCTTAACCAAGAAAA 1200

Db |||||
QY 429 LeuTrpIleProGlyPheAlaThrLeuAlaAlaProLeuTyrProLeuThrArgGluGly 448
QY 1201 GGGGAATTCCTCGGCTCTGAGCAGCAGAGGCATTTGATGTATCAAAAAGGCCCTG 1260
Db |||||
QY 449 IleProGluTrpLysGluHisGlnArgAlaPheGluAlaIleLysSerSerLeu 468
QY 1261 CTGAGCCACCTGCTCTGGCCCTCCCTGACGTAACACCCCTTTACCTTTATGTGAT 1320
Db |||||
QY 469 MetThrAlaProAlaLeuAlaLeuProAspLeuThrLysSerPheValLeuTyrValAsp 488
QY 1321 GAGCGTAGGAGTAGCCCGGGAGTTTAAACCCCAAAACCTTAGGACCATGAGAGACCT 1380
Db |||||
QY 489 GluArgAlaGlyIleAlaArgGlyValLeuThrGlnAlaLeuGlyProTrpLysArgPro 508
QY 1381 GTCGCCCTACTGTCAAAAGAGCTCGATCCTGTAGCCAGTGGTTGGCCCATATGCTGAAG 1440
Db |||||
QY 509 ValAlaTyrLeuSerLysLysLeuAspProValAlaSerGlyTrpProThrCysLeuLys 528
QY 1441 GCTATCCGAGTGGCCATCTCGTCAAGAGCCTGACAAATTGACTTTGGGACAGAAT 1500
Db |||||
QY 529 AlaIleAlaValAlaLeuLeuIleLysAspAlaAspLysLeuThrMetGlyGlnGln 548
QY 1501 ATAACTGTAAATAGCCCCCATGTCATTCGAGAAACATCGTTCCGCGAGCCCCAGACCGATGG 1560
Db |||||
QY 549 ValThrValValAlaProHisAlaLeuGluSerIleValArgGlnProProAspArgTrp 568
QY 1561 ATGACCAACGGCCGCTGACCCACTATCAAAAGCCTGCTTCCACAGAGAGGTCACGTTTC 1620
Db |||||
QY 569 MetThrAsnAlaArgMetThrHisTyrGlnSerLeuLeuLeuAsnAspArgValThrPhe 588
QY 1621 GCTCCACGACCGCTCTCAACCTGCCACTCTCTGCTGGAAGACACTGTATGAACACAGTG 1680
Db |||||
QY 589 AlaProProAlaIleLeuAsnProAlaThrLeuLeuProLeuThrAsnAspSerValPro 608
QY 1681 ACTCATGATTCCTCATCACTATTGATTGAGAGACTGGGGTCCGCAAGGACCTTACAGAC 1740
Db |||||
QY 609 ValHisArgCysAlaAspIleLeuAlaGluIleGlyThrArgLysAspLeuThrAsp 628
QY 1741 ATACCGTGAAGTGGAGAGTCTAACTGGTTCTACTGACGGAAGCAGTATGTGGTGAA 1800
Db |||||
QY 629 GlnProTrpProGly---AlaProSerTrpTyrThrAspGlySerSerPheLeuIleGlu 647
QY 1801 GGTAAAGAGTGGCTGGGGCGGTGGTGACGGGACCCGACATCTGGCGCAGAGC 1860
Db |||||
QY 648 GlyLysArgArgAlaGlyAlaValValAspGlyLysLysValIleTrpAlaSerAla 667
QY 1861 CTGCGGGAAGAACTTTCAGCACAAAAGGCTGAGCTCATGGCCCTCACCAAGCTTTGCGG 1920
Db |||||
QY 668 LeuProGluGlyThrSerAlaGlnLysAlaGluLeuIleAlaLeuThrGlnAlaLeuArg 687
QY 1921 CTGCGCGAAGGGAATTCATAAACATTTATACGACAGCAGGTATGCTTTGCGACTGCA 1980
Db |||||
QY 688 GluAlaGluGlyLysIleIleAsnIleTyrThrAspSerArgTyrAlaPheAlaThrAla 707
QY 1981 CACCTACATGGGCGCATCTATAAACAAGGGGTTGCTTACTCAGCAGGAGGGAATA 2040
Db |||||
QY 708 HisIleHisGlyAlaIleTyrArgGlnArgGlyLeuLeuThrSerAlaGlyLysAspIle 727
QY 2041 AAGCAACAAGAGGAAATTCCTAAGCCTATTAGAAGCCGTACATTTTACCAAAAAGGCTAGCT 2100
Db |||||
QY 728 LysAsnLysGluGluIleLeuAlaLeuLeuGluAlaIleHisAlaProLysLysValAla 747
QY 2101 ATTATACACTGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACCATG 2160
Db |||||
QY 748 IleIleHisCysProGlyHisGlnLysGlyGlyAspLeuValAlaLysGlyAsnArgMet 767
QY 2161 GCTCAGCGGTTGCCAAGCAGCCAGCGGTGTAACTTCTGCTCTATATAGAAATG 2220
Db |||||
QY 768 AlaAspSerValAlaLysGlnValAlaGlnGlyAlaMetIleLeuThrGluLysGlyAsn 787
QY 2221 CCC---AAAGCCCAAGACCCAGCAGCATCACCTCTAGAAAGACTGGCAGAGATAAAA 2277
Db |||||

```
Db      788 ProSerLysSerProGlu-----AspGluAsnTyr 797
Qy      2278 AAGATAGACCAAGTTCTCTGAGACTCGGAAGGACCTGCTATACCTCAGATGGGAAGAA 2337
Db      798 AspiLeLysGluLeuPheTyrTrpSerAspProLeuProTyrPhePheGluGlyIle 817
Qy      2338 ATCTGCCCCCAAGAAGGGTTAGATATGTCACAGATACATGCTCTAACCCACCTA 2397
Db      818 AspLeuThrProGluGluGlyIleLysPheValLysGlyLeuHisGlnPheThrHisLeu 837
Qy      2398 GGAATAACACCTGTCAGACAGTGGTCAGAACATCCCTTATCATGTTCTGAGGCTACCA 2457
Db      838 GlyValGluLysMetMetArgLeuIleLysLysSerArgTyrGlnValProAsnLeuLys 857
Qy      2458 GGAGTGGCTGACTCGGTGGTCAAAACATTGTGCTGCCAGCTGGTTAATGCTAACTCT 2517
Db      858 SerValAlaGlnLysIleLeuAsnSerCysLysAlaCysAlaPheThrAsnAlaThrLys 877
Qy      2518 TCCAGATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGGGCTCACTGGGAAGTG 2577
Db      878 ThrTyrLysGluProGlyLysArgGlnArgGlyAspArgProGlyValTyrTrpGluVal 897
Qy      2578 GACTTCTACTGAGGTAAAGCCGGCTAAATACGGAACAAATACCTATTGGTTTGTAGAC 2637
Db      898 AspPheThrGluValLysProGlyMetTyrGlyAsnLysTyrLeuLeuValPheValAsp 917
Qy      2638 ACCTTTTTCAGGATGGTAGAGCTTATCTACTAAGAAAGAGACTTCAACCGCTGGCTG 2697
Db      918 ThrPheSerGlyTrpValGluAlaPheProThrLysThrGluThrAlaGlnIleValAla 937
Qy      2698 AAAAAATACTGGAAGAAATTTTCCAGATTTGGAATACCTAAGGTAATAGGTCAGAC 2757
Db      938 LysLysIlePheGluGluLeuLeuProArgTyrGlyValProLysValIleGlySerAsp 957
Qy      2758 AATGGTCCAGCTTTTGTGGCCAGTAAGTCAGGACTGGCCAGATATTTGGGATTCAT 2817
Db      958 AsnGlyProAlaPheValAlaGlnValSerGlnGlyLeuAlaThrGlnLeuGlyIleAsp 977
Qy      2818 TGGAAACTGCAATTTGTCATACAGACCCCAAGCTCAGACAGGTAGAGAGGATGAATAGA 2877
Db      978 TrpLysLeuHisCysAlaTyrArgProGlnSerSerGlyGlnValGluArgMetAsnArg 997
Qy      2878 ACCATTAAAGAGACCTTACTAATTGACCGCGGAGACTGGCGTTAATGATGGATAGCT 2937
Db      998 ThrLeuLysGluThrLeuThrLysLeuAlaMetGluThrGlyLysAspTrpValAla 1017
Qy      2938 CTCTGCGCTTTGTGCTTTTGTAGGTTAGGACACCCCTGGACAGTTGGCTGACCCCC 2997
Db      1018 LeuLeuProLeuAlaLeuPheArgAlaArgAsnThrProGlyArgPheGlyLeuThrPro 1037
Qy      2998 TATGAATTACTCTACGGGGGACCCCGCCCATTTGGTAGAAATTGCTTCT-----GTACAT 3051
Db      1038 PheGluValLeuTyrGlyGlyProProLeuIleLysAspGlyGlyThrLeuValPro 1057
Qy      3052 AGTGCTGACGTGCTGCTTTCCAGCCTTTGTTCTTAGCTCAAGGCACTTGAGTGGGTG 3111
Db      1058 AspSerGlySerValLeuProSerSerLeuLeuIleHisLeuLysAlaLeuLysValIle 1077
Qy      3112 AGACACGACGGTGGAGGCACTCCGGGAGCCCTACTCAGAGGAGAGACTTCAGATC 3171
Db      1078 ArgThrGlnIleTrpAspGlnLeuLysThrAlaTyrThr---ProGlyThrThrAlaVal 1096
Qy      3172 CCACATCGTTTCAAGTGGGAGATTCACTGTACGTCAGCCACCGCTGAGGAAACCTC 3231
Db      1097 ProHisGluPheGlnValGlyAspGlnValLeuValArgArgHisArgThrGlySerLeu 1116
Qy      3232 GAGACTCGGTGGAGGCCCTTATCTCGTACTTTTGTACCAACCAACCGCTGGAAGTC 3291
Db      1117 GluProArgTrpLysGlyProTyrLeuValLeuLeuThrThrProThrAlaValLysVal 1136
Qy      3292 GAAGGATCTCCACCTGGATCCATGATCCAGCTTAAACCGGGGCCACCTCCCGAT--- 3348
Db      1137 AspGlyIleAlaSerTrpIleHisAlaSerHisValLysArgAlaProSerGlnAspGlu 1156
```

```
Qy      3349 -----TCGGGTGGAAAGCCGAAAGAGCTGAAATCCCTTAAGCTTCGCCTC 3396
Db      1157 GluThrHisGluAspAsnTrpAlaValGluAlaThrAspAsnProLeuLysLeuArgLeu 1176
Qy      3397 CATCGCTGGTTCTCT 3411
Db      1177 LeuArgArgSerPro 1181
RESULT 3
US-09-309-572-13
; Sequence 13, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; EARLIER FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag-pol protein
; US-09-309-572-13
```

```
Alignment Scores:
Pred. No.: 0 Length: 1737
Score: 3797.50 Matches: 719
Percent Similarity: 76.3% Conservative: 165
Best Local Similarity: 62.0% Mismatches: 242
Query Match: 61.4% Indels: 33
DB: 2 Gaps: 12
```

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-09-309-572-13 (1-1737)

```
Qy      4 GGTGCCACAGGGCAACACAGTATCCATGACTACCCGAAAGAACAGTTGCTGGAGTG 63
Db      589 GlyAlaThrGlyGlyLysArgTyrArgTyrThrAspArgLysValHisLeuAlaThr 608
Qy      64 GGAGGGTAACCCACCTGCTTCTGGTCATACCTGAGTGGCCAGCACCCTCTTAGGTAGA 123
Db      609 GlyLysValThrHisSerPheLeuHisValProAspCysProTyrProLeuLeuGlyArg 628
Qy      124 GACTTATTGACCAAGATGGGAGCACAAATTTCTTTTGAA---CAAGGGAAACCCAGAAAGTG 180
Db      629 AspLeuLeuThrLysLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValMet 648
Qy      181 TCTGCAATAACAAACCTATCACTGTGTTGACCTCCAAATTAGATGACGAATATCGACTA 240
Db      649 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu 668
Qy      241 TACTCTCCCTAGTAAGCTGATCAAAATATA---CAATTCTGTTGGAACAGTTTCCC 297
Db      669 HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 688
Qy      298 CAAGCTGGGCAACAAACCCGAGGATGGTTCGGCAAGCAAGTTCCTCCCAAGTTATT 357
Db      689 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleIle 708
Qy      358 CAATCAAGCCAGTCCACACAGTGTCTAGTCAGACAGTACCCCTTTGAGTAAGAAGCT 417
Db      709 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyrProMetSerGlnGluAla 728
Qy      418 CAAGAAGGAATTCCGGCCGATGTCAAAGATTATCAACAGGCGCATCTCTAGTTCTGTCT 477
Db      729 ArgLeuGlyIleLysProHisIleGlnArgLeuAspGlnGlyIleLeuValProCys 748
```


Db 1459 LysProGlyLeuTyxGlyTyrlsTyrlsValPheIleAsePheSerGlyTrp 1478
Qy 2653 GTAGAGGCTTCTCTACTAAGAAAGAGACTTCAACCGGTGGCTTAAATACTGAA 2712
Db 1479 IleGluAlaPheProThrLysLysGluThrAlaLysValValThrLysLysLeuGlu 1498
Qy 2713 GAAATTTTCCAGATTTGGAATACCTAAGTAATAGGTCAGACAAATGGTCCAGCTTTT 2772
Db 1499 GluIlePheProArgPheGlyMetProGlnValLeuGlyThrAspAsnGlyProAlaPhe 1518
Qy 2773 GTTGGCCAGGTAAGTCAGGACTGCGCAAGATATTGGGATTTGATTGGAATCTGATGT 2832
Db 1519 ValSerLysValSerGlnThrValAlaAspLeuGlyLysAspTrpLysLeuHisCys 1538
Qy 2833 GCATACAGACCCCAAGCTCAGGACAGGTCAGAGGATCAATAGAACCAATTAAGAGACC 2892
Db 1539 AlaTyArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThr 1558
Qy 2893 CTTACTAATTCAGCCGAGACTGCGGTGTAATGATGATGATGATGATGATGATGATGAT 2952
Db 1559 LeuThrLysLeuThrLeuAlaThrGlySerArgAspTrpValLeuLeuLeuProLeuAla 1578
Qy 2953 CTTTGTAGGTTAGGAACACCCCTGGACAGTTTGGGCTGACCCCTATGAATTAATCTTAC 3012
Db 1579 LeuTyArgAlaLysGlnThrProGlyProHisGlyLeuThrProTyTrpGluIleLeuTy 1598
Qy 3013 GGGGACCCCTTCTGTTAGAAATTCCTTCTGTACATAGTGTGCTGCTGCTGCTTCC 3072
Db 1599 GlyAlaProProProLeuValAsnPheProAspProAspMetThrArgValThrAsnSer 1618
Qy 3073 CAGCCTTTGCTCTAGGCTCAAGGACTTGAAGTGGGTGAGACAAAGAGCTGGAGGAA 3132
Db 1619 ProSerLeuGlnAlaHisLeuGlnAlaLeuTyrlsLeuValGlnHisGluValTrpArgPro 1638
Qy 3133 CTCGGGAGGCTTACTCAGGAGGAGGACTTGCAG---ATCCACATCGTTTCCAAGTG 3189
Db 1639 LeuAlaAlaTyTrpGlnGlnLeuAspArgProValValProHisProTyTrpArgVal 1658
Qy 3190 GGAGATTCACTAGTTAGACGCCCGTGCAGGAAACCTCGAGACTCGGTGGAAGGCG 3249
Db 1659 GlyAspThrValTrpValArgArgHisGlnThrLysAsnLeuGluProArgTrpLysGly 1678
Qy 3250 CTTATCTCGTACTTTGACACACCAAGCGCTGGAAGTCGAAGGAATCTCCACCTG 3309
Db 1679 ProTyTrpValLeuLeuThrThrProThrAlaLeuLysValAspGlyIleAlaAlaTrp 1698
Qy 3310 ATCATGATCCACGTTAAACCGCGCCACT-----CCGATTG-----GGG 3354
Db 1699 IleHisAlaAlaHisValLysAlaAlaAspProGlyGlyGlyProSerSerArgLeuThr 1718
Qy 3355 TGGAAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGGTGTTCCCT 3411
Db 1719 TrpArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGluAlaPro 1737

RESULT 4

US-09-718-096-13
; Sequence 13, App89763
; Patent No. 6589763
; GENERAL INFORMATION:
; APPLICANT: Von Leer, Meike-Dorothee
; TITLE OF INVENTION: RETROVIRAL HYBRID VECTORS PSEUDOTYPED WITH LCMV
; FILE REFERENCE: 35-195
; CURRENT APPLICATION NUMBER: US/09/718,096
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: DE 19856463.5
; PRIOR FILING DATE: 1998-11-26
; PRIOR APPLICATION NUMBER: EP 99250415.9
; PRIOR FILING DATE: 1999-11-25
; PRIOR APPLICATION NUMBER: US 09/309,572
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13

; LENGTH: 1737
; TYPE: PRT
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag-pol protein
US-09-718-096-13

Alignment Scores:

Score: 0 Length: 1737
Pred. No.: 3797,50 Matches: 719
Percent Similarity: 76.3% Conservative: 165
Best Local Similarity: 62.0% Mismatches: 242
Query Match: 61.4% Indels: 33
DB: 2 Gaps: 12

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-09-718-096-13 (1-1737)

Qy 4 GGTGCCACAGGCAACACAGTATCCATGGACTCCCGAAGAACAGATTGACTTGGAGTG 63
Db 589 GlyAlaThrGlyGlyLysArgTyArgTrpThrAspArgLysValHisLeuAlaThr 608
Qy 64 GGACGGGTAAACCACTCGTTTCTGGTCATACCTGAGTCCCGACGACCCCTTAGGTAGA 123
Db 609 GlyLysValThrHisSerPheLeuHisValProAspCysProTyTrpProLeuGlyArg 628
Qy 124 GACTTATTGACCAAGATGGGAGCACAAATTTCTTTTGA---CAAGGGAAACCAAGATG 180
Db 629 AspLeuLeuThrLysLeuLysAlaGlnIleHisPheGluGlySerGlyAlaGlnValMet 648
Qy 181 TCTGCAATTAACAACTCATCTGTTGACCTCCAAATTTAGATGACGAATATCGACTA 240
Db 649 GlyProMetGlyGlnProLeuGlnValLeuThrLeuAsnIleGluAspGluHisArgLeu 668
Qy 241 TACTCTCCCTAGTAAAGCCTGATCAAAATATA---CAATTCTGGTTGGAACAGTTTCCC 297
Db 669 HisGluThrSerLysGluProAspValSerLeuGlySerThrTrpLeuSerAspPhePro 688
Qy 298 CAAGCCTCGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTTCCCGACAAATTTATT 357
Db 689 GlnAlaTrpAlaGluThrGlyGlyMetGlyLeuAlaValArgGlnAlaProLeuIleLe 708
Qy 358 CAACTGAAGCCAGTGCACACCACTGCTAGTCAGTACAGTACCTCTGAGTAAAGAACT 417
Db 709 ProLeuLysAlaThrSerThrProValSerIleLysGlnTyTrpMetSerGlnGluAla 728
Qy 418 CAAGAAAGAAATTCGGCGCATGTCCAAAGATTAAATCCAAAGGCGATCTTGTCTCTC 477
Db 729 ArgLeuGlyIleLysProHisIleGlnArgLeuLeuAspGlnGlyIleLeuValProCys 748
Qy 478 CAATCTCCCTGGAAATCTCCCTGCTACCGTTAGAAAGCTGGGACTTAATGACTATCGA 537
Db 749 GlnSerProTrpAsnThrProLeuLeuProValLysLysProGlyThrAsnAspTyArg 768
Qy 538 CCAGTACAGACTTGAGAGAGTCAATAAACGGGTGCGAGGATATACACCCAGTCCCG 597
Db 769 ProValGlnAspLeuArgGluValAsnLysArgValGluAspIleHisProThrValPro 788
Qy 598 AACCTTATAACCTCTGTGTGCTCTCCACCCCAACCGAGCTGATATACAGTATTGGAC 657
Db 789 AsnProTyTrpAsnLeuLeuSerGlyLeuProProSerHisGlnTrpTyTrpValLeuAsp 808
Qy 658 TTAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCAACCACTTTTTCCTTC 717
Db 809 LeuLysAspAlaPhePheCysLeuArgLeuHisProThrSerGlnProLeuPheAlaPhe 828
Qy 718 GAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACTGGAGCCCGACCTCCCAA 777
Db 829 GluTrpArgAspProGluMetGlyIleSerGlyGlnLeuThrTrpThrArgLeuProGln 848
Qy 778 GGGTTCAAGAACTCCCGGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTC 837
Db 849 GlyPheLysAsnSerProThrLeuPheAspGluAlaLeuHisArgAspLeuAlaAspPhe 868

Db	1579	LeuTyArgAlaAraGAsnThrProGlyProHisGlyLeuThrProTyArgIuLeLeuTyR	1598
Qy	3013	GGGGGACCCCCCATTTGGTAGAAATGCTTCTACATAGTGTGACGTGCTGCTTTCC	3072
Db	1599	GlyAlaProProLeuValAsnPheProAspProAspMetThrArgValThrAsnSer	1618
Qy	3073	CAGCCTTTGTCTCTAGGCTCAAGGACCTTTGAGTGGGTGAGACAACAGCGGTGAGGCA	3132
Db	1619	ProSerLeuGlnAlaHisLeuGlnAlaLeuTyLeuValGlnHisGluValTArgPro	1638
Qy	3133	CTCGGGAGGCGCTACTCAGGAGGAGGACTTGCAG---ATCCACATCGTTTCCAA	3189
Db	1639	LeuAlaAlaIaTyGlnGluLeuAspArgProValValProHisProTyArgVal	1658
Qy	3190	GGAGATTCACTTACGTTAGACGCCACCGTGCAGGAACCTCGAGACTCGTGAAGG	3249
Db	1659	GlyAspThrValTArgValArgHisGlnThrLysAsnLeuGluProArgTArgGly	1678
Qy	3250	CCATTCTCGTACTTTTGACACACACCGCTGTGAAGTGCAGAACTCCACCTCG	3309
Db	1679	ProTyThrValLeuLeuThrThrProThrAlaLeuLysValAspGlyIleAlaIaT	1698
Qy	3310	ATCCATGCATCCACGTTAAACCGCGCCACCT-----CCCGATTGC-----	3354
Db	1699	IleHisAlaAlaHisValLysAlaAlaAspProGlyGlyGlyProSerSerArgLeuThr	1718
Qy	3355	TGGAAAGCCGAAAGACTGAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCT	3411
Db	1719	TrpArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGluAlaPro	1737

RESULT 5

US-08-329-967-8
; Sequence 8, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Rupprecht, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-Length cDNA Libraries

```
%
% COMPUTER READABLE FORM:
% MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
% COMPUTER: IBM PC compatible
% OPERATING SYSTEM: PC-DOS/MS-DOS
% SOFTWARE: WinPatIn (Genentech)
% CURRENT APPLICATION DATA:
% APPLICATION NUMBER: US/08/929,967
%
```

```

? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Dreger, Ginger R.
? REGISTRATION NUMBER: 33,055
? REFERENCE/DOCKET NUMBER: P103SR1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-3216
? TELEFAX: 650/952-9881
?
```

TOPOLGY:
US-08-929-967-8

Alignment Scores:			
Pred. No.:	0	Length:	1079
Score:	3600.00	Matches:	679
Percent Similarity:	76.7%	Conservative:	156
Best Local Similarity:	62.4%	Mismatches:	222

Query	Match:	58.2%	Indels:	32
DB:		1	Gaps:	11
US-10-723-552-3_COPY_2307_5741 (1-3435) x US-08-929-967-8 (1-1079)				
Qy	211	ACCCCTCAATYTAGATGAGCAATATCGCACTATCTCTCCCTAGTAAAGCGCTGATCAAAAT	270	
Db	1	ThrLeuAsnIleGluAspGluHiAspLeuHiGluThrSerLysGluProAspValSer	20	
Qy	271	ATA---CAATTCCTGTTGGAAACAGTTTCCCAAGCCTCGGCGAGAAACCGCAGGATGGGT	327	
Db	21	LeuGlySerThrTrpLeuSerAspPheProGlnAlaTrpAlaGluThrGlyMetGly	40	
Qy	328	TTGGCAAGACGAGTTCCCCCAAGTATTATTCACCTGAAGCGCAGTGCCACACGATGTCA	387	
Db	41	LeuAlaValArgGlnAlaProLeuIleIleProLeuLysAlaThrSerThrProValSer	60	
Qy	388	GTCAAGACAGTACCCCTTGAGTAAAGAGAGCTCAAGAAGGAATTCGGCCGCATCTCCCAAGA	447	
Db	61	IleLysGlnIyrProMetSerGlnGluAlaGluGlyIleLysProHisIleGlnArg	80	
Qy	448	TTAATCCCAACAGGGCATCTAGTTCTCTGTCCAAATCTCCCTGGAACTCTCCCTGTACCG	507	
Db	81	LeuLeuAspGlnGlnIleLeuValProCysGlnSerProTrpAsnThrProLeuLeuPro	100	
Qy	508	GTTAGAAGCCTGGGACTAATGACTATTCGACAGTACAGGACTTCAGAGAGTCAATATAA	567	
Db	101	ValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLys	120	
Qy	568	CGGTGCAGGATATACACCCAAACAGTCCCGAACCTTATAACCTCTCTGTGCTCTCCCA	627	
Db	121	ArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeuPro	140	
Qy	628	CCCCAACGGAGCTGGTATACAGTATTGGACTTAAAGGATGCCTTCTCTGTGCTGAGATTA	687	
Db	141	ProSerHisGlnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeu	160	
Qy	688	CACCCACTAGCCAAACCACTTTTTCCTTCGAATGGAGAGATCCAGGTACGGGAGAACC	747	
Db	161	HisProThrSerGlnProLeuPheAlaPheGlnTrpArgAspProGluMetGlyIleSer	180	
Qy	748	GGGCAGCTCACCTGGACCCGACTCCCCCAAGGGTTCAAGAACTCCCCACCATCTTTGAC	807	
Db	181	GlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPheAsp	200	
Qy	808	GAGGCCCTACAGAGACCTGGCCCACTTCAGATGCCAACACCTCAGGTGACCTCTCTC	867	
Db	201	GluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuIleLeuLeu	220	
Qy	868	CAGTACGTGATGACCTGCTTCCTGGCGGAGCCACCAACAGGACTGCTTGAAGCGCACG	927	
Db	221	GlnTyrValAspAspLeuLeuLeuAlaThrSerGluLeuAspCysGlnGlnGlyThr	240	
Qy	928	AAGGCCTACTGCTGGAAATTGTCGACCTAGGCTACAGAGCCTCTGTGAAGAGCCCGAC	987	
Db	241	ArgAlaLeuLeuGlnThrLeuGlyAsnLeuGlyTyrArgAlaSerAlaLysLysAlaGln	260	
Qy	988	ATTTGCAGGAGAGAGGTAAATACTCTGGGTACAGTTTGGCGGAGCGGCAGGATGGCTG	104	
Db	261	IleCysGlnLysGlnValLysTyrLeuGlyTyrLeuLeuLysGluGlyGlnArgTrpLeu	280	
Qy	1048	ACGGAGCGCAGGAAGAAAATCTGAGTCCAGATACCGGCCCCCAACACCAAGCAACAAATG	1107	
Db	281	ThrGluAlaArgLysGluThrValMetGlyGlnProThrProLysThrProArgGlnLeu	300	
Qy	1108	AGAGAGTTTTTGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGGCACCTTA	1167	
Db	301	ArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGluMet	320	
Qy	1168	GCAGCCCCCTACCGCTTAACCAAGAAAAGGGGGAATTCCTCTGGCTCTCTGACGAC	1222	
Db	321	AlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAspGln	340	

QY 1285 CTGACGTAACTAACCTTTACCTTTATGTGTGATGAGCGTAAGGAGTAGCCCGG 1341
Db 361 ProAspLeuThrLysProPheGluLeuPheValAspGluLysGlnGlyTyrAlaLys 379

RESULT 8
US-08-979-847B-89
; Sequence 89, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; JOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSER: OLIEF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-Nov-6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 89:
US-08-979-847B-89

Alignment Scores:
Pred. No.: 2.57e-127 Length: 768
Score: 1415.50 Matches: 318
Percent Similarity: 58.9% Conservative: 117
Best Local Similarity: 43.0% Mismatches: 255
Query Match: 22.9% Indels: 49
DB: 2 Gaps: 14

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-08-979-847B-89 (1-768)

QY 4 GTGCGCACAGGGGACACACATATCCATGAGTACC-----CGAAGAACAGTT 51
Db 61 GlyValLeuGlyGln-----ProValThrThrTyrPheSerGlnProLeuSerCys 77

QY 52 GACTTGGAGTGGGACGGTAACCCACTCGTTCTGTGTATCATCTGAGTCCCGACCCC 111
Db 78 AspTrpGlyThrLeuLeuPheSerHisAlaPheLeuIleMetProGluSerProThrPro 97

QY 112 CTCTTAGTAGAGACTTATTGACCAAGATGGGACGACAAATTTCTTTTGAACAAGGAAA 171
Db 98 LeuLeuGlyArgAspIleLeuAlaLysAlaGlyAlaIleIleHisLeuAsnIleGlyLys 117

QY 172 CCAGAAGTGCTGCAAAATAACAACCTATCACTGTGTGGACCCCTCCATTAAGTACGAA 231
Db 118 -----GlyIleProIleCys----- 122

QY 232 TATCGACTATACTCTCCCTAGTAAAGCCTGATCAAAATATACAAATCTGTGTGGACAG 291
Db 123 -----CysProLeuLeuGluGlyIleAsn----- 131

QY 292 TTTCCCAAGCCTGGGCGAGAAACCGCAGGATGGTTTGGCAAGAGTTCCCCCAAA 351
Db 132 ---ProGluValTrpAlaIleGluGlyGlnTyrGlyGlnAlaLysAsnAlaArgProVal 150

QY 352 GTTATTCAACTGAGGCCAGTGCACACACAGTGTCACTGACGACAGTACCCCTTCAAGTAA 411
Db 151 GlnValLysLeuLysAspSerAlaSerPheProTyrGlnArgLysTyrProLeuArgPro 170

QY 412 GAAGCTCAAGAAGGAATTCGGCCGATGTCCTCAAGATTATCCAAACAGGCGATCCTAGTT 471
Db 171 GluAlaLeuGlnGly***GlnLysIleValLysAspLeuLysAlaGlnGlyLeuValLys 190

QY 472 CTGTCCAATCTCCCTGGAACTCTCCCTGCTACCGGTTAGAAAGCCTGGGACTTAATGAC 531
Db 191 ProCysSerSerProCysAsnThrProIleLeuGlyValArgLysPro---AsnGlyGln 209

QY 532 TATCGACCATACAGGACTTGAGAGAGGTCAATAACGGGTGCGAGGATATATACCCCAAC 591
Db 210 TrpArgLeuValGlnAspLeuArgIleAsnGluAlaValPheProLeuTyrProAla 229

QY 592 GTCCCGAACCTTATAACCTCTGTGTGCTCTCCACCCCAACGGAGCTGTGTATACAGTA 651
Db 230 ValSerSerProTyrThrLeuLeuSerLeuIleProGluGlnAlaGluTppPheThrVal 249

QY 652 TTGAGCTTAAAGGATGCTTCTTCTGCTGAGATTACACCCCATAGTACCCCAACCACTTTT 711
Db 250 LeuAspLeuLysAspAlaPhePheCysIleProValArgProAspSerGlnPheLeuPhe 269

QY 712 GCCTTGAATGGAGAGATCCAGGTACGGGAAGAACCGGGAGCTCCTGACCCCGACTG 771
Db 270 AlaPheGlu-----AspPro---LeuAsnProThrSerGlnLeuThrTppThrValLeu 286

QY 772 CCCCAAGGGTTCAAGAACTCCCGACCATCTTTGACGAAGCCCTTACACAGAGACTCGCC 831
Db 287 ProGlnGlyPheArgAspSerProHisLeuPheGlyGlnAlaLeuAlaGlnAspLeuSer 306

QY 832 AACTTCAGGATCCACACACCTCAGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTG 891
Db 307 GlnPhe-----SerTyrLeuAspThrLeuValLeuGlnTyrValAspAspLeuLeu 324

QY 892 GCGGGAGCCACCAACAGGACTGCTTAGAAGGACGAGGACACTACTGTGGATTTGTTCT 951
Db 325 ValAlaArgSerGluThrLeuCysHisGlnAlaThrGlnGlnLeuLeuThrPheLeuThr 344

QY 952 GACCTAGGCTACAGAGCTCTGCTAAGAGGCCAGATTGTCAGGAGAGAGGTAAACATAC 1011
Db 345 ThrCysGlyTyrLysValSerLysProLysAlaArgLeuCysSerGlnGluIleArgTyr 364

QY 1012 TTGGGGTACAGTTTGGGGACGGGACGATGCTGACGGAGGACGAGGACGAGAAAGAACTGTA 1071
Db 365 LeuGlyLeuLysLeuSerLysGlyThrArgAlaLeuSerGluGluArgIleGlnProIle 384

QY 1072 GTCCAGATACCGGCCCAACCAACAGCCAAACAATAGAGAGAGTTTGGGGACAGCTGGA 1131
Db 385 LeuAlaTyrProHisProLysThrLeuLysGlnLeuArgGlyPheLeuGlyIleThrGly 404

QY 1132 TTTTGCAGCTGTGGATCCCGGGTTTGGACCTTAGCAGCCCTTACGACCCCTCTACCCGCTAAC 1191
Db 405 PheCysArgLysGlnIleProArgTyrThrProIleAlaArgProLeuTyrThrLeu 424

QY 1192 AAGAAAAAGGGGAATTCTCC-----TGGGCTCTCTGAGCACCAGAGGCCA 1236


```

486 AlaThrLeuGlnProAlaThrPheLeuPro---AspAenGluGluGlnIleGluHisAen 504
1690 TGCCATCAACTATTGATGAGCAGACTGGGGTCGCGCAAGGACCTTACAGACATACCGCTG 1749
505 CysGlnGlnValIleAlaGlnThrTyrAlaAlaArgGlyAspLeuLeuGluValProLeu 524
1750 ACTGGAGAGTGCTAACTGCTGCTACGACGGAAGAGCTATGTGGTGAAGTAAAGAGG 1809
525 ThrAspProAspLeuAenLeuTyrThrAspGlySerLeuAlaGluLysGlyLeuArg 544
1810 ATGCTGGGGGGGGTGGTGGAGCGGACCGCAGCATCTGGGCGCAGCAGCTCCCGAA 1869
545 LysAlaGlyTyrAlaValIleSerAspAenGlyIleLeuGluSerAsnArgLeuThrPro 564
1870 GGAATCTTCAGCAACAAAGGCTGAGCTCATGGCCCTTCACGCAAGCTTTGGCGCTGCGCAA 1929
565 GlyThrSerAlaHisLeuAlaGluLeuIleAlaLeuThrTrpAlaLeuGluLeuGlyGlu 584
1930 GGGAAATCCATAAATATTATACGACAGCAGGATGCTTGGCTTGGAGCTGCACATACAT 1989
585 GlyLysArgValAlaIleTyrSerAspSerLysTyrAlaTyrLeuValLeuHisAlaHis 604
1990 GGGGCCCATATATAACAAGGGGTGCTTACCTCAGCAGGAGGGAATAAAGACAAA 2049
605 AlaAlaIleTrpArgGluArgGluPheLeuThrSerGluGlyThrProIleAsnHisGln 624
2050 GAGGAATCTTAAAGCTTATTAGAGCGCTACATTTACCAAAAGGCTAGCTTATTATACAC 2109
625 GluAlaIleArgArgLeuLeuLeuAlaValGlnLysProLysGluValAlaValLeuHis 644
2110 TGTCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACACAGATGGCTGACCGG 2169
645 CysGlnGlyHisGlnGluGluGluArgGluIleGluGlyAsnArgGlnAlaAspIle 664
2170 GTTGCCCAAGCAGGAGCC---CAGGGTGTAACTTCTGCTATATAAGAAATGCC 2223
665 GluAlaLysAlaAlaArgGlnAspSerProLeuGluMetLeuIleGluGlyPro 683

RESULT 11
US-08-979-847B-208
; Sequence 208, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-8400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 208:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 683 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 208:
US-08-979-847B-208

Alignment Scores:
Pred. No.: 1,21e-124 Length: 683
Score: 1387.50 Matches: 311
Percent Similarity: 59.4% Conservative: 116
Best Local Similarity: 43.3% Mismatches: 249
Query Match: 22.4% Indels: 43
DB: 2 Gaps: 13

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-08-979-847B-208 (1-683)

QY 88 GTCATCTAGTGGTCCGAGCAGCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCA 147
Db 1 IleMetProGluSerProThrProLeuLeuGlyArgAspIleLeuAlaLysAlaGlyAla 20
QY 148 CAAATTTCTTTTGAACAAGGGAACCAAGAGTGTCTGCAATAACAACTATCCTCCTGTG 207
Db 21 IleIleHisLeuAsnIleGlyLys-----GlyIleProIleCys--- 33
QY 208 TTGACCTCCAAATTAGATGAGAAATATCGACTATATCTCTCCCTAGTAAAGCTGATCAA 267
Db 34 -----CysProLeuLeuGluGluGlyIle 41
QY 268 AATATACAATTCTGTTGGAACAGTTTCCCAAGCCCTGGGCGAGAAACCGCAGGATGGGT 327
Db 42 Asn-----ProGluValTrpAlaIleGluGlyGlnTyrGly 53
QY 328 TTGGCAAGCAAGTTTCCCAAGTATTCAACTGAAGCCAGTGCACACACAGCTGTCA 387
Db 54 GlnAlaLysAsnAlaArgProValGlnValLysLeuLysAspSerAlaSerPheProTyr 73
QY 388 GTCAGACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCGCATGTCCAAAGA 447
Db 74 GlnArgLysTyrProLeuArgProGluAlaLeuGlnGly**GlnLysIleValLysAsp 93
QY 448 TTAATCCACAGGGCATCTAGTTCCTGTCCTCAATCTCCCTGGAATCTCCCTGCTACCG 507
Db 94 LeuLysAlaGlnGlyLeuValLysProCysSerSerProCysAsnThrProIleLeuGly 113
QY 508 GTTAGAAGCTGGGACTAATGACTATCGACACAGTACAGGACTTGAGAGAGGTCAATAAA 567
Db 114 ValArgLysPro---AsnGlyGlnTrpArgLeuValGlnAspLeuArgIleIleAsnGlu 132
QY 568 CGGGTGAGGATATACACCAACAGTCCCGAACCCTTATACCTCTTGTGTCTCTCCCA 627
Db 133 AlaValPheProLeuTyrProAlaValSerSerProTyrThrLeuLeuSerLeuIlePro 152
QY 628 CCCCACAGGAGCTGTATACAGTATTGGACTTAAAGATGCTCTCTCTGCTGAGATTA 687
Db 153 GluGluAlaGluTrpPheThrValLeuAspLeuLysAspAlaPhePheCysIleProVal 172
QY 688 CACCCCACTAGCCAAACACATTTTTCCTTCGAAATGGAGAGATCCAGTACGGGAAGAAC 747
Db 173 ArgProAspSerGlnPheLeuPheAlaPheGlu-----AspPro---LeuAsnProThr 189
QY 748 GGCAGCTCACCTGGACCCGACTGCCCAAGGGTTCAAGAGGTTCAAGAACTCCCGACCATCTTTGAC 807
Db -----

```

Db 190 SerGlnLeuThrTrpThrValLeuProGlnGlyPheArgAspSerProHisLeuPheGly 209
QY 808 GAAGCCCTACAGAGACTGCGCCAACTTCAGGATCCAAACCCCTCAGGTGACCTCCTC 867
Db 210 GlnAlaLeuAlaGlnAspLeuSer-----GlnProSerTyrLeuAspThrLeuValLeu 227
QY 868 CAGTACGTGGATGACCTCTCTGCGGGAGCCACCAACAGGACTGCTAGAAAGGCACG 927
Db 228 GlnTyrValAspAspLeuLeuValAlaArgSerGluThrLeuCysHisGlnAlaThr 247
QY 928 AAGGCACTACTGCTGGAATTGCTGACCTAGGCTACAGAGCTCTGCTAAAGAGCCGAC 987
Db 248 GlnGluLeuLeuIlePheLeuThrThrCysGlyTyrLysValSerLysProLysAlaArg 267
QY 988 ATTTCAGGAGAGAGTAACTACTTGGGTACAGTTTGGGGAGCGGACGACGATGGCTG 1047
Db 268 LeuCysSerGlnGluIleArgTyrLeuGlyLeuSerLysGlyThrArgAlaLeu 287
QY 1048 ACGGAGGCACGGAAGAACTGTAGTCCAGATACCGGCCCAACACACAGCCCAACAATG 1107
Db 288 SerGluGluArgIleGlnProIleLeuAlaTyrProHisProLysThrLeuLysGlnLeu 307
QY 1108 AGAGAGTTTTGGGACAGCTGGATTGTCAGACTGTGGATCCCGGGTGTTCGACCTTA 1167
Db 308 ArgGlyPheLeuGlyIleThrGlyPheCysArgLysGlnIleProArgTyrThrProile 327
QY 1168 GCAGCCCACTCTACCCGCTAACCAAGAAAGGGGAATTCCTCC----- 1212
Db 328 AlaArgProLeuTyrThrLeuIleArgGluThrGlnLysAlaAsnThrTyrLeuValArg 347
QY 1213 TGGGCTCTGAGCACCAGAGCATTTGATGCTATCAAAAGGCCCTGCTGAGCGCACCT 1272
Db 348 TrpThrProThr---GluValAlaPheGlnAlaLeuLysAlaLeuThrGlnAlaPro 366
QY 1273 GCTCTGGCCCTCCCTGACGTAACTAAACCTTTTACCTTTATGTGGATGAGCGTAAGGGA 1332
Db 367 ValPheSerLeuPro---ThrGlyGlnAspPheSerLeuTyrAlaThrGluLysThrGly 385
QY 1333 GTAGCCCGGGAGTTTAAACCAACCTAGACCATGAGAGACCTGTGCGCTACCTG 1392
Db 386 IleAlaLeuGlyValLeuThrGlnValSerGlyMetSerLeuGlnProValValTyrLeu 405
QY 1393 TCAAGAAAGCTCGATTCGTGACCGAGTGTGGCCCATATGCTGAAGGCTATCGACGT 1452
Db 406 SerLysGluIleAspValValAlaLysGlyTrpProHisCysLeuTrpValMetAlaAla 425
QY 1453 GTGGCCATCTGCTCAAGGACGCTGACAAATGACTTTGGGACAGAAATATAACTGTATA 1512
Db 426 ValAlaValLeuValSerGluAlaValLysIleIleGlnGlyArgAspLeuThrValTrp 445
QY 1513 GCCCCCATGATGGAGAAATGCTTTCGGCAGCCCGACCGATGATGACCAACGCC 1572
Db 446 ThrSerHisAspValAsnGlyIleLeuThrAlaLysGlyAspLeuTrpLeuSerAspAsn 465
QY 1573 CGCATGCCACTATCAAGCTGCTTCTCACAGAGGGTCT---ACGTCGCTCCACCA 1629
Db 466 HisLeuLeuAsnTyrGlnAlaLeuLeuLeuGluProValLeuArgLeuArgThrCys 485
QY 1630 GCGCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGTAGTACCACTGACTCATGAT 1689
Db 486 AlaThrLeuGlnProAlaThrPheLeuPro---AspAsnGluGluLysIleGluHisAsn 504
QY 1690 TGCCATCAACTATGATTGAGAGACTGGGGTCCGCAAGGACCTTACAGACATCCGCTG 1749
Db 505 CysGlnGlnValIleAlaGlnThrTyrTrpAlaAlaArgGlyAspLeuLeuValProLeu 524
QY 1750 ACTGGAGAGTGTCAACCTGTTCTACTGACGGAACGACTGTGTGTGGAAGGTAAAGG 1809
Db 525 ThrAspProAspLeuAsnLeuTyrThrAspGlySerLeuAlaGluLysGlyLeuArg 544
QY 1810 ATGGCTGGGGCGGTGGTGGACCGGACCCGACGATCTGGGCCAGCGCTGCGGAA 1869
Db 545 LysAlaGlyTyrAlaValIleSerAspAsnGlyIleLeuLeuGluSerAsnArgLeuThrPro 564

QY 1870 GGAACCTTCAGCACAAAGGCTGAGCTCATGSCCTCAGCAAGCTTTGCGCTGCGGAA 1929
Db 565 GlyThrSerAlaHisLeuAlaGluLeuIleAlaLeuThrTrpAlaLeuGluLeuGlyGlu 584
QY 1930 GGGAAATCCATAAATATATATACGACAGCAGGATGCTTTGCGACTGCGACACGTACAT 1989
Db 585 GlyLysArgValAsnIleTyrSerAspSerLysTyrAlaTyrLeuValLeuHisAlaHis 604
QY 1990 GGGGCCATCTATAAACAAGGGGTTGCTTACCTCAGCAGGAGGGAATAAAGAACAA 2049
Db 605 AlaAlaIleTrpArgGluArgPheLeuThrSerGluGlyThrProIleAsnHisGln 624
QY 2050 GAGGAATCTTAACCTTAGAAGCGCTACATTTACCAAAAAGGCTAGCTATTATATACAC 2109
Db 625 GluAlaIleArgArgLeuLeuAlaValGlnLysProLysGluValAlaLeuHis 644
QY 2110 TGTCTCTGACATCAGAAAGCTAAAGATCTATATCCAGAGAAACACGATGGCTGACCGG 2169
Db 645 CysGlnGlyHisGlnGluGluArgGluLeuGluArgGluLeuGluAsnArgGlnAlaAspIle 664
QY 2170 GTTCCCAAGCAGCGAGCC---CAGGGTGTAACTTCTGCTCTATATAAGAAATGCC 2223
Db 665 GluAlaLysLysAlaAlaArgGlnAspSerProLeuGluMetLeuIleGluGlyPro 683

RESULT 12

US-08-979-847B-210
; Sequence 210, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:

APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE, PLC

STREET: P.O. BOX 19928

STATE: VA

CITY: ALEXANDRIA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-No. 6582703-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BERRIDGE, WILLIAM P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 39046A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEO ID NO: 210:

SEQUENCE CHARACTERISTICS:

LENGTH: 683 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 210:

US-08-979-847B-210

Alignment Scores:

Pred. No.:	1,21e-124	Length:	683
Score:	1387,50	Matches:	311
Percent Similarity:	59.4%	Conservative:	116
Best Local Similarity:	43.3%	Mismatches:	249
Query Match:	22.4%	Indels:	43
DB:	2	Gaps:	13

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-08-979-847B-210 (1-683)

```
Qy 88 GTCATACCTGAGTCCGACAGCCCTCTTAGGTAGAGACTTATTGACCAAGATGGAGCA 147
Db 1 IleMetProGluSerProThrProLeuLeuGlyArgAspIleLeuAlaLysAlaGlyAla 20
Qy 148 CAATTTCTTTTGAACAGGGAACACAGAGTGTCTGCAATATACAAACCTATCACTGTG 207
Db 21 IleIleHisLeuAsnIleGlyLys-----GlyIleProIleCys--- 33
Qy 208 TTGACCCCTCCAATTAGATGACGAATATCGATATACTCTCCCTAGTAAAGCCTGATCAA 267
Db 34 -----CysProLeuLeuGluGluGlyIle 41
Qy 268 AATATACAAATTCTGTTGGAACAGTTTCCCAAGCCTGGGACAGAAACCGCAGGGATGGT 327
Db 42 Asn-----ProGluValTrpAlaIleGluGlyGlnTrpGly 53
Qy 328 TTGCAAGCAGTCTCCCGCACAGTATTTCACATGAAGCCAGTGCACACAGTGTCA 387
Db 54 GlnAlaLysAsnAlaArgProValGlnValLysLeuLysAspSerAlaSerPheProTrp 73
Qy 388 GTACAGACAGTACCCCTTGAGTAAAGAGTCAAGAAGAAATTCGGCCGATGTCCAAAGA 447
Db 74 GlnArgLysTrpProLeuArgProGluAlaLeuGlnGly***GlnLysIleValLysAsp 93
Qy 448 TTAATCAACAGGCGATCTAGTTCCTGTCATTCCTGGAATCTCCCTGGAATCTCCCTG 507
Db 94 LeuLysAlaGlnGlyLeuValLysProCysSerSerProCysAsnThrProIleLeuGly 113
Qy 508 GTTAGAAGCCTGGGATATGATATGACATCGACAGTACAGGACTTGAGAGAGTCAATAAA 567
Db 114 ValArgLysPro---AsnGlyGlnTrpArgLeuValGlnAspLeuArgIleIleAsnGlu 132
Qy 568 CGGTCGAGGATATACACCAACAGTCCCGAACCTTATACCTCTTGTGTGCTCTCCCA 627
Db 133 AlaValPheProLeuTrpAlaValSerSerProTrpTrpThrLeuLeuSerLeuIlePro 152
Qy 628 CCCCAACGGAGCTGGTATACAGTATTGGACTTAAAGGATGCTTCTTCTGCTGAGATTA 687
Db 153 GluGluAlaGluTrpPheThrValLeuAspLeuLysAspAlaPhePheCysIleProVal 172
Qy 688 CACCCCACTAGCCAAACCACTTTTTCCTTGAATGGAGAGATNCOAGTACGGGAAGAAC 747
Db 173 ArgProAspSerGlnPheLeuPheAlaPheGlu-----AspPro---LeuAsnProThr 189
Qy 748 GGGCAGCTCACCTGACCCGACTGCCCCAAGGGTTCAAGAACTCCCGACCATCTTTGAC 807
Db 190 SerGlnLeuThrTrpThrValLeuProGlnGlyPheArgAspSerProHisLeuPheGly 209
Qy 808 GAAGCCCTCACACAGAGACTGGCCAACTTCAGGATCCAAACCCCTCAGGTGACCCCTCCT 867
Db 210 GlnAlaLeuAlaGlnAspLeuSer-----GlnProSerTrpLeuAspThrLeuValLeu 227
Qy 868 CAGTACGTGATGACTGCTCTTGGCGGAGCCACCAACAGGACTGCTTAGAAGGCAGC 927
Db 228 GlnTrpValAspAspLeuLeuValAlaArgSerGluThrLeuCysHisGlnAlaThr 247
Qy 928 AAGGCACCTACTGCTGGAATGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAGGCCAG 987
Db 248 GlnGluLeuLeuIlePheLeuThrThrCysGlyTrpLysValSerLysProLysAlaArg 267
```

```
Qy 988 ATTTCAGGAGAGAGAGTAAACATCTTGGGTTACAGTTTGGGGACGGCAGCGATGGCTG 1047
Db 268 LeuCysSerGlnGluIleArgTrpLeuGlyLeuLysLeuSerLysGlyThrArgAlaLeu 287
Qy 1048 ACGGAGCAGCAGGAAGAAATCTAGTCCAGATACCGGCCCAACACAGCCAAACAATG 1107
Db 288 SerGluGluArgIleGlnProIleLeuAlaTrpProHisProLysThrLeuLysGlnLeu 307
Qy 1108 AGAGAGTTTTCGGGACAGCTGGATTTTGACAGCTGTGGATCCCGGGTTTCGACCTT 1167
Db 308 ArgGlyPheLeuGlyIleThrGlyPheCysArgLysGlnIleProArgTrpThrProIle 327
Qy 1168 GCAGCCCACTTACCTACCGCTAACCAAGAAAGGGAATTCCTC----- 1212
Db 328 AlaArgProLeuTrpThrLeuIleArgGluThrGlnLysAlaAsnThrTrpLeuValArg 347
Qy 1213 TGGGCTCTCAGCAGCAGAGGCAATTTGATGCTATCAAAAAGCCCTCTGAGCGCACCT 1272
Db 348 TrpThrProThr---GluValAlaPheGlnAlaLeuLysLysAlaLeuThrGlnAlaPro 366
Qy 1273 GCTCTGCGCTCCTCGTACGTAACCTTAAACCTTTTACCTTTTATGTGGATGACGTAAGGA 1332
Db 367 ValPheSerLeuPro---ThrGlyGlnAspPheSerLeuTrpAlaThrGluLysThrGly 385
Qy 1333 GTAGCCCGGGAGTTTAAACCAACCTTAGACCATAGGAGACCTGTGCTACCTG 1392
Db 386 IleAlaLeuGlyValLeuThrGlnValSerGlyMetSerLeuGlnProValValTrpLeu 405
Qy 1393 TCAAGAGCTCGATCTCTGAGCAGTGTGGCCATATGCTTGAAGGCTATCGCAGCT 1452
Db 406 SerLysGlnIleAspValValAlaLysGlyTrpProHisCysLeuTrpValMetAlaAla 425
Qy 1453 GTGGCCATCTGGTCAAGGACCGTGCACAAATTTGACTTTGGGACAGAAATAAATGTAATA 1512
Db 426 ValAlaValLeuValSerGluAlaValLysIleIleGlnGlyArgAspLeuThrValTrp 445
Qy 1513 GCCCCCATCATTTGGAGAACATCGTTTCGGAGCCCCCAGACCGATGATGACCAAGCC 1572
Db 446 ThrSerHisAspValAsnGlyIleLeuThrAlaLysGlyAspLeuTrpLeuSerAspAsn 465
Qy 1573 CGATGACCCACTATCAAGCCTGCTCTCACAGAGAGGGTC---ACGTTGCTCCACCA 1629
Db 466 HisLeuLeuAsnTrpGlnAlaLeuLeuGluGluProValLeuArgLeuArgThrCys 485
Qy 1630 GCGCTCTCAACCTGCTGCTTCTGCTTGAAGAGACTGATGAACAGTACTCATCATG 1689
Db 486 AlaThrLeuGlnProAlaThrPheLeuPro---AspAsnGluGluLysIleGluHisAsn 504
Qy 1690 TGGCATCAACTATTGATTGAGGAGACTGGGGTCGCAAGGACCTTACAGACATACCCGCTG 1749
Db 505 CysGlnGlnValIleAlaGlnThrTrpAlaAlaArgGlyAspLeuLeuGluValProLeu 524
Qy 1750 ACTGGAGAGTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1809
Db 525 ThrAspProAspLeuAsnLeuTrpThrAspGlySerSerLeuAlaGluLysGlyLeuArg 544
Qy 1810 ATGCTGGGGCGCGCTGTGGAGCGGACCCGACAGCATCTGGGCGACAGCTCCCGCGAA 1869
Db 545 LysAlaGlyTrpAlaValIleSerAspAsnGlyIleLeuGluSerAsnArgLeuThrPro 564
Qy 1870 GGAACCTTACACAAAGGCTGAGCTCATGCCCTCTACGCAAGCTTTGCGGCTGCCCGAA 1929
Db 565 GlyThrSerAlaHisLeuAlaGluLeuIleAlaLeuThrTrpAlaLeuGluGlyGlu 584
Qy 1930 GGAATATCCATAAATTTATACGACAGCAGTATGCTTTCGCGCTGACACAGTACAT 1989
Db 585 GlyLysArgValAsnIleTrpSerAspSerLysTrpAlaTrpLeuValLeuHisAlaHis 604
Qy 1990 GGGGCCATCTATAAACAAGGGGTTGCTTACTCTCAGCAGGAGGGGAAATAAAGAACAAA 2049
Db 605 AlaAlaIleTrpArgGluArgGluPheLeuThrSerGluGlyThrProIleAsnHisGln 624
Qy 2050 GAGGAAATCTTAAGCCTATTAGAAGCGGTACATTTTACCAAAAAAGGCTAGCTATTATAC 2109
```



```
Db 370 SerGlyMetSerLeuGlnProValValTyrLeuSerLysGluLeuAspValValAlaLys 389
Qy 1420 GTTGGCCCATATGCTGGAAGCTATCGCAGCTGTGGCCATACCTGTCGAAGAGCTGAC 1479
Db 390 GlyTrpProHisCysLeuTrpValMetAlaAlaValAlaValLeuValSerGluAlaVal 409
Qy 1480 AAATGTGACTTTGGGACAAATATACTTAATAGCCCCCATGATGGAGAAACATGTT 1539
Db 410 LysIleIleGlnGlyArgAspLeuThrValTrpThrSerHisAspValAsnGlyIleLeu 429
Qy 1540 CGGCAGCCCCCAGCCGATGACCAACGCCCGCATGACCCACTATCAAAAGCTGTT 1599
Db 430 ThrAlaLysGlyAspLeuTrpLeuSerAspAsnHisLeuAsnTyrGlnAlaLeuLeu 449
Qy 1600 CTCACAGAGAGGTC---ACGTGCTCCACAGCGGCTCTCAACCTGCGCACCTCTCTG 1656
Db 450 LeuGluGluProValLeuArgLeuArgThrCysAlaThrLeuLysProAlaThrPheLeu 469
Qy 1657 CTTGAGAGACTGATGAACAGTGAATCATGATGATGCCATCAACTATTGATTGAGAGACT 1716
Db 470 Pro---AspAsnGluGluLysIleGluHisAsnCysGlnGlnValIleAlaGlnThrTyr 488
Qy 1717 GGGGTCCGACAGGACCTTACAGACATACCGCTGACTGGAGAGTGCCTAACTGGTTCCT 1776
Db 489 AlaAlaArgGlyAspLeuLeuGluValProLeuThrAspProAspLeuAsnLeuTyrThr 508
Qy 1777 GACGGAAGCAGCTATGTGGTGAAGTAAGAGATGCTGGCGCGCGGTGGTGGACGGG 1836
Db 509 AspGlySerSerLeuAlaGluLysGlyLeuArgLysAlaGlyTyrAlaValIleSerAsp 528
Qy 1837 ACCGCGACGATCTGGCGCAGCAGCTGCGGGAAGAACTTCAGCACAAAAGCTGAGCTC 1896
Db 529 AsnGlyIleLeuGluSerAsnArgLeuThrProGlyThrSerAlaHisLeuAlaGluLeu 548
Qy 1897 ATGCGCCTCAGCAAGCTTTCGGCTGCGCCAGGGAATCCATAACATTTATACGAC 1956
Db 549 IleAlaLeuThrTrpAlaLeuGluLeuGlyGlyLysArgValAsnIleTyrSerAsp 568
Qy 1957 AGCAGGTATGCTTTGGGACTGACACAGCTATCGGGGCATCTATAACAAGAGGGGTG 2016
Db 569 SerLysTyrAlaTyrLeuValLeuHisAlaHisAlaIleTrpArgGluArgGluPhe 588
Qy 2017 CTTACCTCAGCAGGAGGGAATTAAGAACAAAGAGGAATTTAAGCTTATTAGAAGCC 2076
Db 589 LeuThrSerGluGlyThrProIleAsnHisGlnGluAlaIleArgLeuLeuLeuAla 608
Qy 2077 GTACATTTACAAAAGCTAGCTATTATACACTGCTCTGACATCAGAAGCTAAAGAT 2136
Db 609 ValGlnLysProLysGluValAlaValLeuHisCysGlnGlyHisGlnGluGluGlu 628
Qy 2137 CTCATATCCAGAGGAACACAGATGGCTGACCGGTTGCCAAGCAGGAGGCCAG 2190
Db 629 ArgGluIleGluGlyAsnArgGlnAlaAspIleGluAlaLysLysAlaAlaArg 646
```

RESULT 14

```
US-09-120-653D-5
; Sequence 5, Application US/09120653D
; Patent No. 6365727
; GENERAL INFORMATION:
; APPLICANT: YOON, JI-WON
; APPLICANT: JUN, HEE-SOOK
; APPLICANT: PARK, HAE-JOON
; APPLICANT: AHN, JONG-SEONG
; APPLICANT: HA, YOUNG-JU
; APPLICANT: CHUNG, SOO-IL
; TITLE OF INVENTION: DNA and peptides of a diabetes-specific endogenous retrovirus
; FILE REFERENCE: 98-338
; CURRENT APPLICATION NUMBER: US/09/120,653D
; CURRENT FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: KR 98-10108
; PRIOR FILING DATE: 1998-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: KopatentIn 1.71
```

```
; SEQ ID NO 5
; LENGTH: 768
; TYPE: PRT
; ORGANISM: DIABETES-SPECIFIC RETROVIRUS ENDOGENOUS ERV-9
; FEATURE:
; NAME/KEY: NON_CONS
; LOCATION: (46)
; OTHER INFORMATION: any one of amino acids
; FEATURE:
; NAME/KEY: NON_CONS
; LOCATION: (193)
; OTHER INFORMATION: any one of amino acids
; FEATURE:
; NAME/KEY: NON_CONS
; LOCATION: (322)
; OTHER INFORMATION: any one of amino acids
; FEATURE:
; NAME/KEY: NON_CONS
; LOCATION: (374)
; OTHER INFORMATION: any one of amino acids
; FEATURE:
; NAME/KEY: NON_CONS
; LOCATION: (526)
; OTHER INFORMATION: any one of amino acids
; FEATURE:
; NAME/KEY: NON_CONS
; LOCATION: (586)
; OTHER INFORMATION: any one of amino acids
US-09-120-653D-5
```

Alignment Scores:

Pred. No.:	5,65e-110	Length:	768
Score:	1236.00	Matches:	305
Percent Similarity:	55.2%	Conservative:	106
Best Local Similarity:	41.0%	Mismatches:	270
Query Match:	20.0%	Indels:	63
DB:	2	Gaps:	18

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-09-120-653D-5 (1-768)

```
Qy 73 ACCACCTGTTTCTGGTCATACCTGAGTCCGACGACCCCTCTTAGGTAGAGACTTATTG 132
Db 80 SerHisAlaPheLeuValIleProGluSerProThrProLeuLeuGlyArgAspIleLeu 99
Qy 133 ACCAAGATGGAGCACAAATTTCTTTTGAACAGGAAACCCAGAAAGTGTCTGCAGAAATAAC 192
Db 100 AlaLysAlaGlyAlaIleThr-TyrMetAsnLysArgAsnLys----- 113
Qy 193 AAACCTATCACTGTGTTGACCCCTCCAAATTAGATGACGAATATCGACTATCTCTCCCTA 252
Db 114 ----LeuThrIleCysCysPro-----LeuLeuGluArgGluSerThrLeu---- 127
Qy 253 GTAAGCCTGATCAAAATATACAAATTCTGTTGGACAGTTTCCCCAAGCCTGGGCGAGAA 312
Db 128 -----LysSerGlyHisTrpLysAspAsnLeuGluGlyGlnLy 140
Qy 313 ACCGACGGGATGGTTTGGCAAGCAAGTTCCCCCAAGTTATTCAACTGAAGGCCAGT 372
Db 140 bMetProAlaGlnSerLysSerGlyLysArgProHisPhe----- 154
Qy 373 GCCACACCACTGTCAGTCAGACAGTACCCCTTGGTAAAGAAGCTCAAGAAGGAATTCGG 432
Db 155 -----SerTyrGln--ArgGlnTyrProLeuArgProGluAlaHisLysGlyLeuGln 171
Qy 433 CCGCATGTCOAAGATTAAATCCAAACAGGGCATCTCTAGTCTCTGTCCTTCCCTGGAAT 492
Db 172 AspIleAlaGluHisLeuIleAlaGlnGlyLeuValArgLysCysSerSerPro---Asn 190
Qy 493 ACTCCC---CTGCTACCGGTTAGAAAGCCCTGGGCACTAATGACTATCGACCACTACAGAC 549
Db 191 ThrPro***IleLeuGlyValGlnLysPro---AsnGlyGlnTrpArgLeuValGlnAsp 209
Qy 550 TTGAGAGAGTCAATAAACCAGGTCAGGATATACACCAACAGTCGCCGAAACCTTATAAC 609
```


CITY: Abbott Park
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60064
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
 MEDIUM TYPE: storage
 COMPUTER: IBM
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WordPerfect 5.1
 CURRENT APPLICATION DATA: US/08/259,451
 APPLICATION NUMBER: 436
 FILING DATE: 20-JUN-1994
 CLASSIFICATION: 436
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/086,415
 FILING DATE: 01-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Daniel W. Collins
 REGISTRATION NUMBER: 31,912
 REFERENCE/DOCKET NUMBER: 5381.US.P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (708) 937-6365
 TELEFAX: (708) 938-2623
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 917 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 US-08-259-451-11

Alignment Scores:

Pred. No.: 2,51e-51 Length: 917
 Score: 629.50 Matches: 274
 Percent Similarity: 36.6% Conservative: 115
 Best Local Similarity: 25.8% Mismatches: 425
 Query Match: 10.2% Indels: 249
 DB: 2 Gaps: 38

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-08-259-451-11 (1-917)

QY 346 CCACAGTATTCACTGAGGCGCAGTGCACACAGTGTGCTGAGTGCAGTACCCCTTG 405
 Db |||||
 QY 38 ProArgThrProSerProThrAlaProGlyGlyProIleSerPheLys----- 53
 QY 406 AGTAAAGAGCTCAAGAGGAATTCGGCGCATGTCCAAAGATTATCCACAGGGCATC 455
 Db |||||
 QY 54 ---ProGluArgLeuGlnAlaLeuAsnAspLeuValSerLysAlaLeuGluAlaGlyHis 72
 QY 466 CTAGTCTCTCCCAATCTCCCTGGAATACTCCCTGCTACCGGTTAGAAAGCTGGGACT 525
 Db |||||
 QY 73 IleGluProTyrSerGlyProGlyAsnAsnProValPheProValLysProPro---Asn 91
 QY 526 AATGACTATGACACAGTACAGGACTTCAGAGAGGTCAATAAACGGGTGCAGGATATACAC 585
 Db |||||
 QY 92 GlyLysTrpArgPheIleHisAspLeuArgAlaThrAsnAlaIleAlaThrLeuThr 111
 QY 586 CCAACAGTCCGACCCCTTATACCTCTTGTTGCTCTCCACCCCAACCGAGCTGGTAT 645
 Db |||||
 QY 112 SerProSerProGlyProProAsp---LeuThrSerLeuProThrAlaLeuProHisLeu 130
 QY 646 ACAGTATTGCACTTAAGAGTATGCTCTCTGCTGAGATTACACCCACTAGCCACCA 705
 Db |||||
 QY 131 GlnThrIleAspLeuThrAspAlaPhePheGlnIleProLeuProLysGlnPheGlnPro 150
 QY 706 CTTTTCCTTCGAATGAGAGATCCA---GGTACGGGAAGAACCGGGCAGCTCACCTGG 762
 Db |||||
 QY 151 TyrPheAlaPheThrIleProGlnProCysAsnTyrGlyProGlyThrArgTyrAlaTip 170
 QY 763 ACCGAGTCCCAAGGTTCAAGAACTCCCGACCATCTTTTCAGAACGCCCTTACACAGA 822
 Db |||||
 QY 171 ThrValLeuProGlnGlyPheLysAsnSerProThrLeuPheGluGlnLeuAlaAla 190

QY 823 GACCTGGCCAACTTTCAGGATCCAAACACCTCAGTGACCTCTCCAGTACGTGATGAC 882
 Db |||||
 QY 191 ValLeuAsnProMetArgLysMetPheProThrIleValGlnTyrMetAspAsp 210
 QY 883 CTGCTTCTGGGGAGCCACCAACAGGACTGCTTAGAAGCAGACGACGACTACTGTGTG 942
 Db |||||
 QY 211 IleLeuLeuAlaSerProThrAsnLysGlnGlnLeuSerGlnLeuThrLeuGln 230
 QY 943 GAATTGTCTGACCTAGGCTACAGAGCTCTGCTAAGAAGGCCCAAGATTTCAGGAGAG 1002
 Db |||||
 QY 231 AlaLeuThrThrHisGlyLeuProIleSerGlnGlnLysThrGlnArgThrProGlyGln 250
 QY 1003 GTAACATACCTTGGGTACAGTTTTCGGGACGGGACGCGATGGCTGACGAGGACCGAAG 1062
 Db |||||
 QY 251 IleArgPheLeuGlyGlnValIleSerProAsnHis-----IleThrTyrGluSer 267
 QY 1063 AAAACTGTACTCAGATACCGGCCCAACACAGCACAACAAATGAGAGAGTTCCTGGG 1122
 Db |||||
 QY 268 ThrProAlaIleProIleLysSerGlnTrpThrLeuThrGluLeuGlnValIleLeuGly 287
 QY 1123 ACAGCTGGATTTTCAGACTGTGATCCCG-----GGTTTTCGACCTTACCA 1170
 Db |||||
 QY 288 GluIleGlnTrpValSerLysGlyThrProIleLeuArgLysHisLeuGlnSerLeuTyr 307
 QY 1171 GCCCCACTCTACCCGCTAACCAAGAAAGGGAATTCCTGGGCTCTTGAGCACCAG 1230
 Db |||||
 QY 308 SerAlaLeuHisGlyTyrArgAspProArgAlaCysIleThrLeuThrProGlnGlnLeu 327
 QY 1231 AAGCATTTGATGCTATCAAAAGGCCCTCTGAGCGCACCTGCTGGGCCCTCCCTGAC 1290
 Db |||||
 QY 328 HisAlaLeuHisAlaIleGlnGlnAlaLeuGlnHisAsnCysArgGlyArgLeu---Asp 346
 QY 1291 GTAACATAACCTTTTACCCTTTATGTGGATGAGCGTAAGGAGTAGCCCGGGAGTTTA 1350
 Db |||||
 QY 347 ProThrLeuProLeuLeuGlyLeuIleSerLeuSerThrSerGlyThrThrSerValIle 366
 QY 1351 ACCCAACCTTAGGACCATCGAGAAGACCTGTGCGCTTACCTGTCAAGAAAGCTCGATCCT 1410
 Db |||||
 QY 367 PheGlnProLysGlnAsnTrp-----ProLeuAlaTrpLeuHisThrProHisProPro 384
 QY 1411 GTAGCC---AGTGGTGGCCCATATGCTGAAGGCTATCCGAGCTGTGGCCATCTGCTGC 1467
 Db |||||
 QY 385 ThrSerLeuCysProTrpGlyHisLeuLeu-----AlaCysThrIle 398
 QY 1468 AAGGACGCTGACAAATTCCTTTGGGACAGAAATTAACGTGTAATA----- 1512
 Db |||||
 QY 399 LeuThrLeuAspLysTyrThrLeu---GlnHisTyrGlyLeuLeuCysGlnSerPheHis 417
 QY 1513 -----GCCCCCAT----- 1521
 Db |||||
 QY 418 HisAsnMetSerLysGlnAlaLeuCysAspPheLeuArgAsnSerProHisProSerVal 437
 QY 1522 -----GCATTGGAGAACATCGTTCCGGACAGCCCGACACCGA 1557
 Db |||||
 QY 438 GlyIleLeuIleHisMetGlyArgPheHisAsnLeuGlySerGlnProSerGlyPro 457
 QY 1558 TGGATGACCAACCCCGCATGACCCACTATCAAGCTGCTTCTCACAGAGAGGCTCAGC 1617
 Db |||||
 QY 458 TrpLysThr-----LeuLeuHisLeuProThrLeuLeu----- 469
 QY 1618 TTGCTCTCCACGACCGCTCTCAACCTGCGCATCTTCTGCTGAGAGACTGATGAACCA 1677
 Db |||||
 QY 469 -----GlnGluProArgLeuLeuArgProIlePheThrLeuSerPro 482
 QY 1678 GTGACTCATGATTCATCAACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACA 1737
 Db |||||
 QY 483 ValValLeuAspThrAlaProCysLeu----- 491
 QY 1738 GACATACCGCTGACTGGAGAGTGTAACTGTTTCACTGACCGAAGCAGCTATGTGGTG 1797
 Db |||||
 QY 492 -----PheSerAspGlySerPro----- 497

QY	1798	GAGGTTAAGAGGATGGCTGGGGCGCGGTGGTGGACGGGACCCGCACGATCTGG-----	1855
Db	498	-----GlnLysAlaAlaTyrValLeuTrpAspGlnThrIleLeuGlnGln512	
QY	1852	---GCCAGCAGCCTGCCG-----GAAGGAACCTTCAGACAAAAGGCTGAGCTCATGGCC1902	
Db	513	AspIleThrProLeuPProHisGluThrAsnSerAlaGlnLysGlyGluLeuLeuAla532	
QY	1903	CTCAGCGAAGCTTTGGCGCTGGCCGAA-----GGGAAATCCATAAACATTTATACGGCAGC1959	
Db	533	LeuIleTyrGlyLeuArgAlaAlaLysProTrpProSerLeuAsnIlePheLeuAspSer552	
QY	1960	AGGTATGCCCTTTGGCACTGCACAC-----GTACATGGGCGCATCTATAAACAAAGGGGG2013	
Db	553	LysTyrLeuIleLysTyrLeuHisSerLeuAlaIleGlyAlaPhe-----567	
QY	2014	TTGCTTACCTCAGCAGGGAGGAAATAAGACAAAGAGGANAATCTTAAGCTTATTAGAA2073	
Db	568	LeuGlyThrSerAlaHisGlnThrLeuGln-----AlaAlaLeuPro581	
QY	2074	GCCGTACATTTACCAAAAGCGTAGCTATTATACACTGCTCCTGGACATCAGAAAGCTAAA2133	
Db	582	ProLeuLeuGlnGlyLysThrIleTyrLeuHisValargSerHisThrAsnLeuPro601	
QY	2134	GATCTCATATCCAGAGAAACAGATGGCTGACCGGTTGCCAAGCAGGCAGCCAGGGT2193	
Db	602	AspProIleSerThrPheAsnGluTyrThrAspSerLeuIle-----615	
QY	2194	GTTAACCTTCTGCCTATAATAGANAATGCCCAAGCCCAAGCCAGAACCCAGACGACAGTACAC2253	
Db	616	-----ValAlaProLeuValProLeu-----622	
QY	2254	CTGAAGACTGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTTCGGGAAGGGAC2313	
Db	623	-----ThrProGlnGly---626	
QY	2314	TGCTATACCTCAGATGGGAAGGAAATCTGTGCCCAAGAAAGGGTTAGAAATATGTCAA2373	
Db	626	-----626	
QY	2374	CAGATACATGCTCTAACCCACCTTAGGAACATAAACCTCGACGAGCTGGTCAGACACATCC2433	
Db	627	---LeuHisGlyLeuThrHisCysAsnGlnArgAlaLeuValSerPheGlyAlaThr---644	
QY	2434	CCTTATCATGTTCTGAGGCTACCAAGGAGTGGCTGACTCGGTGGTCAACCATTTGTGTGCC2493	
Db	645	-----ProlysgluAlaLysSerLeuValGlnThrCysHisThr657	
QY	2494	TGCCAGCTGGTTAATGTCTATCTCTCCAGAAATGCCTCCAGGGAAGAGACTAAGGGGAAGC2553	
Db	658	CysGlnIleIleAsnSerGln---HisHisMetProGlnGlyHisIleArgAspGlyLeu676	
QY	2554	CACCCAGGCGTCACTGGGAAGTGGACTTCACTGAGTAAAGCCGGCTAAATACGGGAAC2613	
Db	677	LeuProAsnHisIleTrpGlnGlyAspValThrHisTyrLysTyrLysLysTyr-----694	
QY	2614	AAATACCTATTG---GTTTTGTAGACACCTTTTCAGATGGGTAGAGGCTTATCTCTACT2670	
Db	695	LysTyrCysLeuHisValTrpValAspThrPheSerGlyAlaValSerValSerCysLys714	
QY	2671	AAGAAAGAGACTTCAACCGTGGCTGAAATAAATACTGGAAGAGAAATTTTTTCCAAGATTT2730	
Db	715	LysLysGluThrSerCysGluThrIleSerAlaPheLeuGlnAlaIle---SerLeuLeu733	
QY	2731	GGAATACCTAAGGTAAATAGGCTCAGACAATGGTCCAGTCTTGTGGTCCCGAGGTAAAGTCAG2790	
Db	734	GlyLysProLeuHisIleAsnThrAspAsnGlyProAlaPheLeuSerGlnGluPheGln753	
QY	2791	GGACTGGCCCAAGATATTGGGGATTGATTGGAAATCGCATTTGTCATACAGACCCCAAGC2850	
Db	754	GluPheCysThrSerTyrHisIleGlyHisSerThrHisIleProTyrAsnProThrSer773	
QY	2851	TCAGACAGGTTAGAGAGGATGAATAGAACCAATTAAAGAGACCCCTTACTATAA-----2901	

Db	774	SerGlyLeuValGluArgThrAsnGlyIleIleLysAsnLeuLeuAsnLysTyrLeuLeu	793
Qy	2902	-----TTGACCGCGAGACTGGCGTTAAATGATTGGATAGCTCTCTCGCCCTTT	2949
Db	794	AspCysProAsnLeuProLeuAspAsnAlaIleAsnLys-----AlaLeuTrpThrLeu	811
Qy	2950	GTGCTTTTGGGGTTAGGAACACCCCTGGACACTTTGGGTGACCCCTATGAATTACTC	3009
Db	812	AsnGlnLeuAsnValMetAsnProSerGlyLys-----ThrArgTrpGlnIleHis	828
Qy	3010	TACGGGGAGCCCCCCCATTGGTAGAAATGTCTTGTATACATAGTGTGACGTGCTGCTT	3069
Db	829	His-----ProProLeuProProIleProGluAlaSerThrProProLysProPro	846
Qy	3070	TCCAGACCTTTGTCTTAGGCTCAAGGCACCTTGAGTGGGTGAGACACAGCGCTGGAGG	3129
Db	847	SerLysTrpPheTyrTyrLysLeuProGlyLeu-----	857
Qy	3130	CAACTCCGGGAGCCCTACTCAGGAGGAGAGACTTCGAGATCCCATCGTTTCCAAGTG	3189
Db	857	-----	857
Qy	3190	GGAGATTCAGTCTAGTTAGACGCCCGTCAGGAAACCTCGAGACTCGGTGAGAGGC	3249
Db	858	-----ThrAsnGlnArgTrpLysGly	864
Qy	3250	CCTTATCTCGTACTTTTGACACCAACCGGCT-----GTGAAGTCGAAGAAATCTCC	3303
Db	865	ProLeuGlnSerLeuGlnGluAlaAlaGlyAlaAlaLeuLeuSerIleAspGlyPhePro	884
Qy	3304	ACCTGGATCCATGCATCCACAGTTAAACCGGCG-----CCACCTCCGATTCGGGGTGG	3357
Db	885	ArgTrpIleProTrpArgPheLeuLysAlaAlaCysProArgProAspAlaSerGlu	904
Qy	3358	AAAGCCGAA	3366
Db	905	ProAlaGlu	907

Search completed: February 14, 2006, 16:17:52
Job time : 226.767 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:07:51 ; Search time 86.3713 Seconds
(without alignments)
3323.427 Million cell updates/sec

Title: US-10-723-552-3_COPY_2307_5741

Perfect score: 6183

Sequence: 1 ATGGGTGCCAGGCGCAACA.....CTGTCAATAACCTCTCAGAC 3435

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10723552/runat_14022006_125148_13205/app.query.fasta_1
-DB=Published Applications_AA_Main -QFWT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsm62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN 1.1 602 runat_14022006_125148_13205 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -BSPLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3808.5	61.6	1199	4	US-10-677-558-2
2	2491	40.3	678	5	US-10-827-498-2
3	2487.5	40.2	716	3	US-08-845-157-2
4	2487.5	40.2	716	4	US-10-661-819-2
5	2007	32.5	471	4	US-10-448-871A-31
6	1867	30.2	471	4	US-10-448-871A-29
7	1862	30.1	471	4	US-10-448-871A-27
8	1850.5	29.9	470	4	US-10-448-871A-28
9	1846	29.9	471	4	US-10-448-871A-30
10	1794	29.0	1438	5	US-10-450-763-42377
11	1784	28.9	473	4	US-10-448-871A-25

12	1768.5	28.6	472	4	US-10-448-871A-26
13	1431.5	23.2	1267	5	US-10-450-763-33781
14	1415.5	22.9	768	2	US-08-979-847-89
15	1415.5	22.9	768	4	US-10-114-104-89
16	1408.5	22.8	817	5	US-10-450-763-40498
17	1406.5	22.7	896	5	US-10-450-763-34798
18	1390.5	22.5	683	2	US-08-979-847-198
19	1390.5	22.5	683	2	US-08-979-847-200
20	1390.5	22.5	683	4	US-10-114-104-198
21	1390.5	22.5	683	4	US-10-114-104-200
22	1387.5	22.4	683	2	US-08-979-847-208
23	1387.5	22.4	683	2	US-08-979-847-210
24	1387.5	22.4	683	4	US-10-114-104-208
25	1387.5	22.4	683	4	US-10-114-104-210
26	1363.5	22.1	821	5	US-10-450-763-52800
27	1345.5	21.8	872	5	US-10-450-763-42647
28	1345.5	21.8	1288	5	US-10-450-763-34793
29	1345.5	21.8	1288	5	US-10-450-763-42636
30	1331.5	21.5	654	2	US-08-979-847-91
31	1331.5	21.5	654	4	US-10-114-104-91
32	1330.5	21.5	1067	5	US-10-450-763-38408
33	1280	20.7	805	5	US-10-450-763-47721
34	1268.5	20.5	997	5	US-10-450-763-39301
35	1248.5	20.2	1253	5	US-10-450-763-57733
36	1208	19.5	694	5	US-10-450-763-37899
37	1181	19.1	1577	5	US-10-450-763-36834
38	1181	19.1	1577	5	US-10-450-763-57482
39	1101.5	17.8	930	5	US-10-450-763-40953
40	1101.5	17.8	1142	5	US-10-450-763-35673
41	1093.5	17.7	1001	5	US-10-450-763-47711
42	1047	16.9	614	5	US-10-450-763-39777
43	1047	16.9	614	5	US-10-450-763-50364
44	1042	16.9	552	5	US-10-450-763-39780
45	1042	16.9	552	5	US-10-450-763-42631

ALIGNMENTS

RESULT 1

US-10-677-558-2
; Sequence 2, Application US/10677558
; Publication No. US20040096972A1
; GENERAL INFORMATION:
; APPLICANT: AUDIT, Muziel
; APPLICANT: COSSET, Francois-Loic
; TITLE OF INVENTION: CHIMERIC PLASMID COMPRISING A REPLICATIVE RETROVIRAL GENOME AND
; FILE REFERENCE: 1759.135
; CURRENT APPLICATION NUMBER: US/10/677,558
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/FR02/03934
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: FR 0114976
; PRIOR FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1199
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: product of pol gene of pAM plasmid
US-10-677-558-2

Alignment Scores: 6.12e-300 Length: 1199
Score: 3808.50 Matches: 722
Best Local Similarity: 76.2% Conservatives: 160
Best Local Similarity: 62.3% Mismatches: 245
Query Match: 61.6% Indels: 31
DB: 4 Gaps: 11

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-677-558-2 (1-1199)

Db 770 ArgMetAlaAspGlnAlaAlaArgGluValAlaThrArgGluThrProGlyThrSerThr 789
Qy 2203 CTGCCTATAATAGAAATGCCCAAGCCCAAGACCCAGACAGACAGTACACCTCAGAACAGC 2262
Db 790 LeuLeuIleGluAsnSerThrProTyrThrHisGluHisPheHisTyrThrValThrAsp 809
Qy 2263 TGGCAGAGATATAAAGATAGACAGCTCTCTGAGACTCCGGAAGGAGACCTGCTATACC 2322
Db 810 ThrLysAspLeuThrLysLeu-----GlyAlaThrTyrAsp 821
Qy 2323 TCA-----GATGGGAAGGAATCTCTGCCCCACCAAGAGGGTTA 2361
Db 822 SerAlaLysLysTyrTyrValTyrGlnGlyLysProValMetProAspGlnPheThrPhe 841
Qy 2362 GAATATGTCCACAGATACATCGTCTAACCCACCTAGGAACCTAAACACCTCGCAGCGTTG 2421
Db 842 GluLeuLeuAspPheLeuHisGlnLeuThrHisLeuSerPheSerLysThrLysAlaLeu 861
Qy 2422 GTCAGA-----ACATCCCTTATCATGTTCTCGAGGCTTACAGAGTGCGTACGCGTG 2475
Db 862 LeuGluArgSerProSerProTyrTyrMetLeuAsnArgAspArgThrLeuLysAsnIle 881
Qy 2476 GTCAAACATCTGTGCTGCCAGCTGTTATCTATCTCTCCAGATCCCTCCAGG 2535
Db 882 ThrGluThrCysLysAlaCysAlaGlnValAsnAlaSerLysSerAlaValLysGlnGly 901
Qy 2536 AAGAGACTAAGGGAAGCACCCAGCGCTCACTGGGAAGTGGACTTCACTCAGGTTAAG 2595
Db 902 ThrArgValAArgGlyHisArgProGlyThrHisTyrGluIleAspPheThrGluValLys 921
Qy 2596 CCGCTAATACGGAACAAATACCTATTGTTTGTAGACACCTTTTCAGGATGGGTA 2655
Db 922 ProGlyLeuTyrGlyTyrLysTyrLeuLeuValPheValAspThrPheSerGlyTyrIle 941
Qy 2656 GAGCTTATCTACTTAAGAAAGACTTCAACCGTGGTGGCTAAATAAATACTCGAAGAA 2715
Db 942 GluAlaPheProThrLysLysGlnThrAlaLysValValThrLysLysLeuGluGlu 961
Qy 2716 ATTTTCCAGATTGGAATACCTAAGGTAATAGGTCAGACAAATGGTCCAGCTTTGTT 2775
Db 962 IlePheProArgPheGlyMetProGlnValLeuGlyThrAspAsnGlyProAlaPheVal 981
Qy 2776 GCCAGGTAGTCAAGGACTCGCCCAAGATATGGGATGATGGAAATCGATGTGCA 2835
Db 982 SerLysValSerGlnThrValAlaAspLeuLeuGlyIleAspTyrLysLeuHisCysAla 1001
Qy 2836 TACAGACCCCAAGCTCAGGACAGTACAGAGGATGAATGAACCATTAAGACACCTT 2895
Db 1002 TyrArgProGlnSerSerGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeu 1021
Qy 2896 ACTAAATTGACCGCGGAGACTGGCGTTAATGATTGGATAGCTCTCTGCCCTTTGTGCTT 2955
Db 1022 ThrLysLeuThrLeuAlaThrGlySerArgAspTyrValLeuLeuLeuProLeuAlaLeu 1041
Qy 2956 TTTAGGTTAGGAACACCCCTGGACAGTTGGCTGACCCCTATGAATTAATCTACGGG 3015
Db 1042 TyrArgAlaArgAsnThrProGlyProHisGlyLeuThrProTyrGluIleLeuTyrGly 1061
Qy 3016 GGACCCCTCCATTTGGTAGAATTTGCTCTCTACATAGTGTGAGCTGCTCTTCCAG 3075
Db 1062 AlaProProLeuValAsnPheProAspProAspMetThrArgValThrAsnSerPro 1081
Qy 3076 CCTTTGTTCTTAGGCTCAAGGCACTTGAGTGGGTGAGAACACGAGCGTGGAGCAACTC 3135
Db 1082 SerLeuGlnAlaHisLeuGlnAlaLeuTyrLeuValGlnHisGluValTyrArgProLeu 1101
Qy 3136 CGGAGGCGCTACTCAGGAGGAGAGACTTGCAG---ATCCACATCGTTTCCAAAGTGGGA 3192
Db 1102 AlaAlaAlaTyrGlnGluGlnLeuAspArgProValValProHisLysProTyrArgValGly 1121
Qy 3193 GATTCAGCTTACGTTAGCCCGCTGTCAGGAGAACCTCGAGACTCGGTGGAAGGCGCT 3252
Db 1122 AspThrValTyrValArgArgHisGlnThrLysAsnLeuGluProArgTyrLysGlyPro 1141

Qy 3253 TATCTCGTACTTTTGACCAACCGGCTGTGAAGTCGAAGGAATCTCCACCTGGATC 3312
Db 1142 TyrThrValLeuLeuThrThrProThrAlaLeuLysValAspGlyIleAlaAlaTyrIle 1161
Qy 3313 CATGATCCCAAGTAAACCGCGCCACCT-----CCGATTCGGGG-----TGG 3357
Db 1162 HisAlaAlaHisValLysAlaAlaAspThrGluSerGlyProSerSerGlyArgThrTyr 1181
Qy 3358 AAAGCCGAAGAGACTGAAATCCCTTAAGCTTCGCTCCATCCGCTGTTTCCT 3411
Db 1182 ArgValGlnArgSerGlnAsnProLeuLysIleArgLeuThrArgGlySerPro 1199
RESULT 2
US-10-827-498-2
; Sequence 2, Application US/10827498
; Publication No. US20050232934A1
; GENERAL INFORMATION:
; APPLICANT: Ambion, Inc.
; APPLICANT: Chen, Lianling
; APPLICANT: Setterquist, Robert
; APPLICANT: Latham, Gary
; TITLE OF INVENTION: Recombinant Reverse Transcriptases
; FILE REFERENCE: AMBI:1001
; CURRENT APPLICATION NUMBER: US/10/827,498
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Moloney murine sarcoma virus
US-10-827-498-2
Alignment Scores:
Pred. No.: 5,37e-193 Length: 678
Score: 2491.00 Matches: 462
Percent Similarity: 82.6% Conservative: 95
Best Local Similarity: 68.5% Mismatches: 107
Query Match: 40.3% Indels: 10
Dbs: 5 Gaps: 4
US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-827-498-2 (1-678)
Qy 208 TTGACCTCCCAATAGATAGCATATATCGACTATATCTCCCTAGTAAAGCCTGATCAA 267
Db 1 MetThrLeuAsnIleGluAspGluTyrArgLeuHisGluThrSerLysGluProAspVal 20
Qy 268 AATATA---CAATTCTGGTTGGACAGTTTCCCAAGCCTGGGAGAAACCGCAGGGATG 324
Db 21 SerLeuGlySerThrTyrLeuSerAspPheProGlnAlaTyrAlaGluThrGlyMet 40
Qy 325 GCTTTGGCAAGCAAGTTCCCCCAAGTATTCACTGAAGCCAGCCAGTGCACACCATG 384
Db 41 GlyLeuAlaValArgGlnAlaProLeuIleIleProLeuLysAlaThrSerThrProVal 60
Qy 385 TCAGTCAGACAGTACCCCTTGTAGTAAAGAAGCTCAAGAAGAAATTCGCGCCATGTCCAA 444
Db 61 SerIleLysGlnTyrProMetSerGlnGluAlaArgLeuGlyIleLysProHisIleGln 80
Qy 445 AGATTAATCAACAGGCGCATCTAGTTCTCTGTCATCTCCCTGGAATACTCCCTGCTGTA 504
Db 81 ArgLeuLeuAspGlnGlyIleLeuValProCysGlnSerProTyrAsnThrProLeuLeu 100
Qy 505 CCGTTAGAAAGCCTGGGACTAATGATATCGACAGTACAGGCTTGAGAGAGTCAAT 564
Db 101 ProValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsn 120
Qy 565 AAACGGGTGAGGATATACACCAACAGTCCCAAGCTTATACCTTATACCTCTTGTGTGCTCTC 624
Db 121 LysArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeu 140
Qy 625 CCACCCCAACGAGCTGGTATACAGTATTGGACTTAAAGGATGCCTTTCTTCCTGCTGAGA 684

```
Db 141 ProProSerHisGlnTrpTyrThrValLeuAspLeuLysAspAlaTyrPheCysLeuArg 160
|||||
695 TTAACCCCACTAGCAACCACTTTTTCGCTTGAATGGAGATCCAGGTACGGGAAGA 744
|||||
Db 161 LeuHisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGluMetGlyIle 180
|||||
745 ACCGGCAGCTACCTGACCGCAGCTGCCCAAGGGTTTCAAGACTCCCGCCACCATCTTT 804
|||||
Db 181 SerGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPhe 200
|||||
805 GACGAAGCCCTACACAGAGACTGCGCAACTTCAGGATCCCAACCCCTCAGGTGACCCCTC 864
|||||
Db 201 AspGluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuIleLeu 220
|||||
865 CTCAGTACGTGATGACTGCTCTTGGCGGAGCCCAACACAGGACTGCTTGAAGGC 924
|||||
Db 221 LeuGlnTyrValAspAspLeuLeuAlaAlaThrSerGluLeuAspCysGlnGlnGly 240
|||||
925 ACGAAGGCACCTACTCTGGAATTGCTGACCTAGGCTACAGAGCTCTGCTAAGAGGCC 984
|||||
Db 241 ThrArgAlaLeuLeuGlnThrLeuGlyAsnLeuGlyTyrArgAlaSerAlaLysLysAla 260
|||||
985 CAGATTTCGAGAGAGAGTACATACTTGGGGTACAGTTTGGCGGACCGGACGAGTGG 1044
|||||
Db 261 GlnIleCysGlnLysGlnValLysTyrLeuGlyTyrLeuLeuLysGluGlyGlnArgTrp 280
|||||
1045 CTGACGAGGCACCGGAAGAACTAGTACGACATACCGGCCCAACACACAGCCCAACAA 1104
|||||
Db 281 LeuThrGluAlaArgLysGluThrValMetGlyGlnProThrProLysThrProArgGln 300
|||||
1105 ATGAGAGATTTTGGGACAGCTGGATTTGCAGACTGCGATCCCGGGTTCGAC 1164
|||||
Db 301 LeuArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGlu 320
|||||
1165 TTAGCAGCCCACTTACCGCTAACCAAGAAAGGGGAATCTCTCGGCTCCTGAC 1224
|||||
Db 321 MetAlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAsp 340
|||||
1225 CACCAAGAGGCATTTGATGCTATCAAAAGCCCTGCTGAGCGCACCTGCTGCGCCCTC 1284
|||||
Db 341 GlnGlnLysAlaTyrGlnGluIleLysGlnAlaLeuLeuThrAlaProAlaLeuGlyLeu 360
|||||
1285 CTGACGTAACCTAACCTTTTACCTTTATGCGATGAGCGTAAAGGAGTAGCCCGGGA 1344
|||||
Db 361 ProAspLeuThrLysProPheGluLeuPheValAspGluLysGlnGlyTyrAlaLysGly 380
|||||
1345 GTTTTAAACCCCAACCTAGGACCATGAGAGACCTGTCGCTACCTGTCAAGAGAGCTC 1404
|||||
Db 381 ValLeuThrGlnLysLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeu 400
|||||
1405 GATCCTGTAGCCAGTGTGGCCCAATATGCTGAAGGCTATGCGAGCTGTGGCCATCTG 1464
|||||
Db 401 AspProValAlaAlaGlyTrpProProCysLeuArgMetValAlaAlaIleAlaValLeu 420
|||||
1465 GTCAGGAGCGCTGACAAATTGATTTGGGACAGATATAACTGTATATAGCCCCCATGCA 1524
|||||
Db 421 ThrLysAspAlaGlyLysLeuThrMetGlyGlnProLeuProLeuValIleAlaProHisAla 440
|||||
1525 TTGAGAGACATCGTTTCGGCAGCCCGCCAGACCGATGGATGACCAACGCCCGATGACCCAC 1584
|||||
Db 441 ValGluAlaLeuValLysGlnProProAspArgTrpLeuSerAsnAlaArgMetThrHis 460
|||||
1585 TATCAAGAGCTGCTTCTC---ACAGAGGGTTCACGTTTCGCTTCCACAGCCGCTCTCAAC 1641
|||||
Db 461 TyrGlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProValValAlaLeuAsn 480
|||||
1642 CTGCACTCTTCTGCTGAAGAGACTGATGACACGAGTCACTCATGATTCATCACTCACTA 1701
|||||
Db 481 ProAlaThrLeuLeuPro---LeuProGluGluGlyLeuGlnHisAsnCysLeuAspIle 499
|||||
1702 TTGATTGAGGAGACTGGGGTCCGCAAGGACCTTTACAGACATACCTGACTGAGAGATG 1761
|||||
```

```
Db 500 LeuAlaGluAlaHisGlyThrArgProAspLeuThrAspGlnProLeuProAspAlaAsp 519
1762 CTAACTCTGTTCACTGACGGAGCAGCTATGTGTGGAGCTAAAGAGTACGCTGGGGCG 1821
|||||
Db 520 HisThrTrpTyrThrAspGlySerSerLeuLeuGlnGluGlyGlnArgLysAlaGlyAla 539
|||||
1822 GCGGTGGTGGACGGGACCCGACCGATCTGGCCAGCAGCCCTCGCGGAAGGAACCTTCAGCA 1881
|||||
Db 540 AlaValThrThrGluThrGluValIleTrpAlaLysAlaLeuProAlaGlyThrSerAla 559
|||||
1882 CAAAGGCTGAGCTCATGCGCCCTCACCAAGCTTTGGCGGTGGCGGAAGGAAATCCATA 1941
|||||
Db 560 GlnArgAlaGluLeuIleAlaLeuThrGlnAlaLeuLysMetAlaGluGlyLysLysLeu 579
|||||
1942 AACATTTATACGACAGCAGGTATGCTTTCGCACTGCACACGTACATGCGGCGCATCTAT 2001
|||||
Db 580 AsnValTyrThrAspSerArgTyrAlaPheAlaThrAlaHisIleHisGlyGluIleTyr 599
|||||
2002 AAACAAAGGGGTGCTTACCTCAGCAGGAGGAGGAAATAAAGAACAAAGAGGAAATCTTA 2061
|||||
Db 600 ArgArgArgGlyLeuLeuThrSerGluGlyLysGluIleLysAsnLysAspGluLeuLeu 619
|||||
2062 AGCCTATTAGAACCGCTACATTTACCAAAAGGCTAGCTATTATACACTGCTCTGGACAT 2121
|||||
Db 620 AlaLeuLysAlaLeuPheLeuProLysArgLeuSerIleIleHisCysProGlyGly 639
|||||
2122 CAGAAAGCTAAGATCTCATATCCAGAGGAACACAGATGGCTGACCGGGTTGCCAAGCAG 2181
|||||
Db 640 GlnLysGlyHisSerAlaGluAlaArgGlyAsnArgMetAlaAspGlnAlaAlaArgLys 659
|||||
2182 GCAGCCCGAGGTGTAACTTCTGCTATATAGAAATGCC 2223
|||||
Db 660 AlaAla-----IleThrGluThrPro 666

RESULT 3
US-09-845-157-2
; Sequence 2, Application US/09845157
; Patent No. US20020090618A1
; GENERAL INFORMATION:
; APPLICANT: Smith, J.
; TITLE OF INVENTION: Thermostable Reverse Transcriptases and Uses Thereof
; FILE REFERENCE: 0942.5040001
; CURRENT APPLICATION NUMBER: US/09/845,157
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: US 60//207,196
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 716
; TYPE: PRT
; ORGANISM: M-MLV reverse transcriptase gene
US-09-845-157-2

Alignment Scores:
Pred. No.: 1,06e-192 Length: 716
Score: 2487.50 Matches: 459
Percent Similarity: 83.5% Conservative: 94
Best Local Similarity: 69.3% Mismatches: 106
Query Match: 40.2% Indels: 3
DB: 3 Gaps: 3

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-09-845-157-2 (1-716)
QY 208 TTGACCTCCAAATTAGATGACGAATATCGACTATATCTCCCTAGTAGTAAGCCTGATCAA 267
|||||
Db 33 MetThrLeuAsnIleGluAspGluTyrArgLeuHisGluThrSerLysGluProAspVal 52
|||||
QY 268 AATATA---CAATTCGTGTGGAACAGTTTCCCAAGCCTGGGAGAAACCGCAGGATG 324
|||||
Db 53 SerLeuGlySerThrTrpLeuSerAspPheProGlnAlaIleTrpAlaGluThrGlyGlyMet 72
|||||
QY 325 GGTTTGGCAAGCAAGTTTCCCGCACAAAGTTATTCAACTGAAGGGCCAGTGCACACCAAGTG 384
|||||
```



```
; FILE REFERENCE: 03316/LH
; CURRENT APPLICATION NUMBER: US/10/448,871A
; CURRENT FILING DATE: 2003-05-29
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 29
; LENGTH: 471
; TYPE: PRT
; ORGANISM: AKV murine leukemia virus
US-10-448-871A-29

Alignment Scores:
Pred. No.: 2,31e-142 Length: 471
Score: 1867.00 Matches: 339
Percent Similarity: 85.6% Conservative: 65
Best Local Similarity: 71.8% Mismatches: 66
Query Match: 30.2% Indels: 2
DB: 4 Gaps: 2

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-448-871A-29 (1-471)

QY 325 GGTGGCAAGCAGTTCCTCCCAAGTATTCACTGAAGGCGCAGTGCACACCAAGTG 384
Db 1 GlyLeuAlaValArgGlnAlaProLeuIlelleProLeuLysAlaThrSerThrProVal 20

QY 385 TCAGTCACAGCAGTACCCCTTGAGTAAGAAGCTCAAGAGGAAATTCGGCGCATGTCCAA 444
Db 21 SerIleLysGlnIleProMetSerGlnGluAlaLysLeuGlyIleLysProHisIleGln 40

QY 445 AGATTAAATCCAAACAGGCGATCTAGTCTCTGTCCAAATCTCCCTGGAATACTCCCTGTCTA 504
Db 41 ArgLeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeuLeu 60

QY 505 CGGTTAGAAAGCCCTGGGACTAATGACTATATCGACAGTACAGGACTTGAGAGAGTCAAT 564
Db 61 ProValLysLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsn 80

QY 565 AAACGGGTGCAGGATATACCCACACAGTCCGACACCTTATACCTCTTGTCCTCTC 624
Db 81 LysArgValGluAspIleHisProThrValProAsnProTyrAsnLeuLeuSerGlyLeu 100

QY 625 CCACCCCAACGGAGCTGTGTATACAGTATTGGAATTAAGAGTGCCTCTTCTGCTCAGAGA 684
Db 101 ProProSerHisArgTyrThrValLeuAspLeuLysAspAlaPheCysLeuArg 120

QY 685 TTACACCCCACTAGCCCAACCACTTTTGGCTTCGAATGGAGAGATCCAGGTACGGGAAGA 744
Db 121 LeuHisProThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyMetGlyIle 140

QY 745 ACCGGGAGCTCACAGACCTGCGCAACTTCAGATCCAAACACCTCCAGTGCACCTC 864
Db 161 AspGluAlaLeuHisArgAspLeuAlaAspPheArgIleGlnHisProAspLeuLeu 180

QY 865 TCCAGTACGTGGATGACTGCTCTGCGGGAGCCACCAACAGCAGTCTGTAGAAGGC 924
Db 181 LeuGlnTyrValAspAspIleLeuSerAlaAlaThrSerGluLeuAspCysGlnGlnGly 200

QY 925 ACGAGGCGACTCTGCGAATGTGTACCTAGGCTACAGGCTCTGCTTAAGAGGCC 984
Db 201 ThrArgAlaLeuLeuLeuLeuLeuGlyAsnLeuGlyTyrArgAlaSerAlaLysLysAla 220

QY 985 CAGATTTGCAGGAGAGAGTAACTACTTGGGGTACAGTTTGGCGGAGCGGCGAGATGG 1044
Db 221 GlnLeuCysGlnLysGlnValLysTyrLeuGlyTyrLeuLeuLysGlnGlyGlnArgTrp 240

QY 1045 CTGACGGAGGCGACGAAGAAGAACTGTGATCCAGATACCGGGCCCCCAACACAGCCCAACA 1104
Db 241 LeuThrGluAlaArgLysGluThrValMetGlyGlnProThrProLysThrProArgGln 260
```

```
QY 1105 ATGAGAGAGTTTTGGGGACAGCTGGATTTTGCAGACTGTGTGATCCCGGGTGTGCGACC 1164
Db 261 LeuArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGlu 280

QY 1165 TTAGACGCCCACTCTTACCCGCTAAACCAAGAAAGGGGAATTCCTCTGGCTCCTGAG 1224
Db 281 MetAlaAlaProLeuTyrProLeuThrLysThrGlyThrLeuPheAsnTrpGlyProAsp 300

QY 1225 CACCAGAGGCATTTGATGCTATCAAAAGCCCTGTGAGCGACCTGTCTGCGCCCTC 1284
Db 301 GlnGlnLysAlaTyrGlnGluIleLysAlaLeuLeuThrAlaProAlaLeuGlyLeu 320

QY 1285 CCTGACGTAACTAAACCTTTTACCCTTTATGTGATGAGCGCTAAGGGAGTAGCCGGGGA 1344
Db 321 ProAspLeuThrLysProPheGluLeuPheValAspGluLysGlnGlyTyrAlaLysGly 340

QY 1345 GTTTTAAACCCAAACCTTAGGACCATGAGGAAGACTGTGCGCTACCTGTCAAAGAGCTC 1404
Db 341 ValLeuThrGlnLysLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeu 360

QY 1405 GATCCTGTAGCCAGTGTGGCCCATATGCTGCTGAAGGCTATCGCAGCTGTGCGCCATCTG 1464
Db 361 AspProValAlaAlaGlyTyrProProCysLeuArgMetValAlaAlaIleAlaValLeu 380

QY 1465 GTCAAGGACGCTGCAAAATGACTTTGGGACAGAAATATAACTGTAATAGCCCCCATGCA 1524
Db 381 ThrLysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAla 400

QY 1525 TTGAGAACATPCGTTCCGACGCCCAACGATGATGATGATGATGATGATGATGATGATGAT 1584
Db 401 ValGluAlaLeuValLysGlnProProAspArgTrpLeuSerAsnAlaArgMetThrHis 420

QY 1585 TATCAAGACCTGCTCTCTC---ACAGAGAGGCTCAGTTCGCTCCACGACCGCTCTCAAC 1641
Db 421 TyrGlnAlaMetLeuLeuAspThrAspArgValGlnPheGlyProValValAlaLeuAsn 440

QY 1642 CCTGCCACTCTCTGCTGCTGAAGAGACTGATGAACACGAGTGCATGATGATGATGATGATGAT 1701
Db 441 ProAlaThrLeuLeuPro---LeuProGluGluGlyAlaProHisAspCysLeuGluIle 459

QY 1702 TTGATGAGGAGACTGGGGTCCGCAAGGACCTTACA 1737
Db 460 LeuAlaGluThrHisGlyThrArgProAspLeuThr 471
```

RESULT 7

```
US-10-448-871A-27
; Sequence 27, Application US/10448871A
; Publication No. US20040093169A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Mizushima, Yoshiyuki
; APPLICANT: Sakaguchi, Kengo
; APPLICANT: Sugawara, Fumio
; TITLE OF INVENTION: A METHOD FOR DESIGNING A MOLECULAR STRUCTURE OF AN INHIBITOR FOR
; FILE REFERENCE: 03316/LH
; CURRENT APPLICATION NUMBER: US/10/448,871A
; CURRENT FILING DATE: 2003-05-29
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 27
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Friend murine leukemia virus
US-10-448-871A-27
```

```
Alignment Scores:
Pred. No.: 5,89e-142 Length: 471
Score: 1862.00 Matches: 341
Percent Similarity: 85.2% Conservative: 61
Best Local Similarity: 72.2% Mismatches: 68
Query Match: 30.1% Indels: 2
DB: 4 Gaps: 2
```

Qy	1405	GATCCTGTAGCAGTGGTTGGCCCATATGCTGAAGGCTATCGCAGCTGTGGCCATACTG	1461
Db	361	AspProValAlaAlaGlyTrpProCysLeuArgMetValAlaAlaIleAlaValLeu	380
Qy	1465	GTCAAGGACCTGCACAAATTGACTTTGGGACAGAAATATAACTGTAAATAGCCCCCATGCA	1524
Db	381	ThrLysAspAlaGlyLysLeuThrMetGlyGlnProLeuValIleLeuAlaProHisAla	400
Qy	1525	TTGAGAACATCGTTTGGCAGAGCCCGAGCCGATGATGACCAACGCCCGCATGACCCAC	1584
Db	401	ValGluAlaLeuValLysGlnProCysArgTrpLeuSerAsnAlaAaGMetThrHis	420
Qy	1585	TATCAAGCCTGCTTCTC---ACAGAGAGGGTCACGTTCCGCTCCACAGCCGCTCTCAAC	1641
Db	421	TyrGlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProIleValAlaLeuAsn	440
Qy	1642	CTGCGCACTCTTCTGCTGAAGAGACTGATGAACACGAGTCACTCATGATGCCATCAACTA	1701
Db	441	ProAlaThrLeuLeuPro---LeuProGluGluGlyLeuGlnHisaspCysLeuGluIle	459
Qy	1702	TTGATTGAGAGACTGGGGTCCGCAAGGACCTTACA	1737
Db	460	LeuAlaGluAlaHisGlyThrArgProAspLeuThr	471
RESULT 8			
US-10-448-871A-28			
; Sequence 28, Application US/10448871A			
; Publication NO. US20040093169A1			
; GENERAL INFORMATION:			
; APPLICANT: Mizushima, Yoshiyuki			
; APPLICANT: Sakaguchi, Kengo			
; APPLICANT: Sugawara, Fumio			
; TITLE OF INVENTION: A METHOD FOR DESIGNING A MOLECULAR STRUCTURE OF AN INHIB			
; TITLE OF INVENTION: AN ENZYME			
; FILE REFERENCE: 03316/LH			
; CURRENT APPLICATION NUMBER: US/10/448, 871A			
; CURRENT FILING DATE: 2003-05-29			
; NUMBER OF SEQ ID NOS: 59			
; SOFTWARE: PatentIn version 2.0			
; SEQ ID NO 28			
; LENGTH: 470			
; TYPE: PRT			
; ORGANISM: Moloney murine leukemia virus			
US-10-448-871A-28			
Alignment Scores:			
Pred. No.:		5.05e-141	Length: 470
Score:		1850.50	Matches: 342
Percent Similarity:		85.2%	Conservative: 60
Best Local Similarity:		72.5%	Mismatches: 67
Query Match:		29.9%	Indels: 3
DB:		4	Gaps: 3
US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-448-871A-28 (1-470)			
Qy	325	GGTTTGGCAAGCAAGTCTCCCCACAGTATTCAACTGAAGCCAGTGCACACCACTG	384
Db	1	GlyLeuAlaValArgGlnAlaProLeuIleIleProLeuLysAlaThrSerThrProVal	20
Qy	385	TCAGTCAGACAGTACCCCTTGAGTAAAGAAGCTCAAGAAGGAATTCGGCGCATGTCCAA	444
Db	21	SerIleLysGlnTyrProMetSerGlnGluAlaAaArgLeuGlyIleLysProHisIleGln	40
Qy	445	AGATTAATCAACAGGGCATCTAGTCTCTGTCCTCAATCTCCCTGGAATACTCCCTTGCTA	504
Db	41	ArgLeuLeuAspGlnGlyIleLeuValProCysGlnSerProTrpAsnThrProLeuLeu	60
Qy	505	CCGTTTAGAAGCCTGGGACTAATGACTATCGACCAAGTACAGGACTTGAGAGAGGTCAAT	564
Db	61	ProValLysLysProGlyThrAsnAspTyrAr-gProValGlnAspLeuArgLeuValAsn	80
Qy	565	AAACGGGTGCAGGATATACACCAACAGTCCCGAACCCCTTATAACCTCTTGTCGTCTTC	624

Db 581 IleProSerProTyThrLeuLeuSerGlnIleProGluGluAlaGluTrpPheThrVal 600
Qy 652 TTGAGCTTAAAGGATGCTTCTTCTGCTGAGATTACACCCGACTAGCCAAACGACTTTT 711
Db 601 LeuAspLeuLysAspAlaPhePheCysIleProValHisProAspSerGlnPheLeuPhe 620
Qy 712 GCCTTCGAATGGAGAGTCCAGGTACGGTAACGGGACCGGCGAGCTCACCTGGACCCGACTG 771
Db 621 AlaPheGlu-----AspPro-----SerAsnProThrSerGlnLeuThrTrpThrValLeu 637
Qy 772 CCCAAGGGTTCAAGAACTCCCGACCATCTTTGACGAAGCCCTACACAGAGACTGGCC 831
Db 638 ProGlnGlyPheArgAspSerProHisLeuPheGlyGlnAlaLeuAlaGlnAspLeuSer 657
Qy 832 AACTTCAGATCCAAACCCCTCAGGTGACCTCCAGTACCCTCCAGTAGCTGGAGTACCTGCTTCTG 891
Db 658 GlnPhe-----SerTyLeuAspThrProValLeuGlnCysMetAspAspLeuLeuLeu 675
Qy 892 GCGGAGGCCCAACACAGGACTGCTTAGAAGGCACGAAGGCACACTACTCTGGAAATTGTCT 951
Db 676 AlaAlaArgSerGluThrLeuCysHisGlnAlaThrGlnAlaLeuLeuAsnPheLeuThr 695
Qy 952 GACCTAGCTACAGAGCTCTCTCTAAGAAAGCCAGATTTCGAGGAGAGAGGTAAACATAC 1011
Db 696 ThrCysGlyTyrlsValSerLysProLysAlaGlnLeuCysSerGlnGlnValLysCys 715
Qy 1012 TTGGGTACAGTTTCGGGACGGGAGCGATGGCTGACGGAGGCACGGAAGAAACTGTA 1071
Db 716 LeuGlyLeuLysLeuSerLysValThrArgAlaLeuSerGluArgIleGlnProIle 735
Qy 1072 GTCCAGATACGGCCCCCAACACACGCCAAACAAATGACAGAGTTTTCGGGACAGCTGGA 1131
Db 736 LeuAlaTyProTyProLysThrLeuLysGlnLeuArgGlyPheLeuGlyIleThrGly 755
Qy 1132 TTTTGCAGACTGTGATCCCGGGTTTCGACCTTAGCAGCCCCACTCTACCCGCTAAC 1191
Db 756 PheCysArgIleTrpIleProArgTyrlsGlyLysIleAlaArgProLeuTyThrLeuIle 775
Qy 1192 AAAGAAAGGGGAATTCCTC-----TGGGCTCTGACACACGAGGCA 1236
Db 776 LysGluThrGlnLysAlaAsnThrHisLeuValArgTrpThrProGluAlaGluAlaAla 795
Qy 1237 TTGTGATCTATCAAAAGCCCTGCTGAGCGACCTGCTCTGGCCCTCCCTGACGTAAC 1296
Db 796 PheHisAlaLeuLysLysAlaLeuMetGlnAlaProValLeuSerLeuLeu-----ThrGly 814
Qy 1297 AACACCTTTACCTTTATGTGATGAGCGTAAAG-----GGAGTA 1335
Db 815 GlnAspPheSerSerTyThrValThrLysAsnLysGlnThrLysLysLysThr***Ile 834
Qy 1336 GCGCGGGAGTTTAAACCAACCCCTAGGACCATGGAGAACCTGTGCGCTACCTGTCA 1395
Db 835 AlaLeuArgValLeuAlaLeuVal***GlyThrSerLeuGlnProValAlaTyrlsSer 854
Qy 1396 AAGAGGCTCATCTGTAGCCAGTGTGGCCCATATGCTGAAGGCTATCGACGCTGT 1455
Db 855 LysGluIleAspValValAlaLysGlyGlnProHisCysLeuArgValValAlaVal 874
Qy 1456 GCCTACTGGTCAGGAGCGCTGACAAATTTGACTTTGGGACAGAAATATACTGTATAGCC 1515
Db 875 AlaValLeuValSerGluAlaValLysIleIleGlnGlyArgAsnLeuThrValTrpThr 894
Qy 1516 CCCCATGTGGAGAAATCTGTCGGGACGCCGCCGAGCCGATGGATGACCAACCCCGC 1575
Db 895 SerHisAspValAsnGlyIleLeuThrAlaLysGlyAspLeuTrpLeuSerAspAsnHis 914
Qy 1576 ATGACCCACTATCAAGCTGCTCTCTCACA---GAGAGGCTCACGTTCTGCTCCACGCC 1632
Db 915 LeuLeuLysTyrlsGlnAlaLeuLeuGluGlyProMetLeuArgLeuCysThrCysAla 934
Qy 1633 GCTCTCAACCCCTGCCACTCTTCTGCTGAAGAGACTGATGACACCACTGATCATGATTC 1692
Db 935 AlaLeuAsnLeuAspThrPheLeuPro---HisAsnGluGluLysIleGluHisAsnCys 953

Qy 1693 CATCACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCCGTGACT 1752
Db 954 GlnGlnValIleAlaGlnThrTyAlaThrArgGlyAspHisLeuGluValProLeuThr 973
Qy 1753 GGAGAAGTGTAACTCGTTTCACTACGAGGAAGCAGCTATGTGTGGAGGTAAGAGGATG 1812
Db 974 AspProAsnProAsnLeuTyThrAspGlyArgSerPheValGluLysGlyLeuGlnLys 993
Qy 1813 GCTGGGCGGCGGTGGTGGAGCCGCGACAGTCTGGCGCAGCAGCTGCGGACGCTCCCGAAGGA 1872
Db 994 ValGlyTyrlsAlaValSerAspAsnGlyIleLeuGluSerAsnProLeuThrProGly 1013
Qy 1873 ACTTCAGCACAAAAGGCTGAGCTCATGCCCTCAGCAAGCTTTTGGCTGGCGCGAAGGG 1932
Db 1014 ThrSerAlaGlnLeuAlaLys-----LeuThrArgAlaLeuGluLeuGlyGly 1030
Qy 1933 AAATCCATAAACATTTATACGACGAGCTATCCCTTTGCGACTGTCACACGTAACATGG 1992
Db 1031 LysArgValAsnIleTyThrAspSerLysTyrlsAlaTyLeuValLeuHisAlaHisAla 1050
Qy 1993 GCCATCTATAAACAAAGGGGTTCCTTACCTCAGCAGGAGGGAATAAAGAACAAAGAG 2052
Db 1051 ValIleTrpArgGluArgGluPheLeuThrSerGluGlyThrProIleLysHisGlnGlu 1070
Qy 2053 GAATTTCTAAGCTATTAGAGCGCTACATTTACCAAAAGCGTAGCTATTATATACACTGT 2112
Db 1071 AlaIleArgGluLeuLeuAlaValGlnLysProLysGluValAlaValLeuHisCys 1090
Qy 2113 CCTGCACATCAGAAAGCTAAAGATCTCATATCCAGAGGAACACGATGGCTGACCGGTT 2172
Db 1091 ArgGlyHisGlnLysGlyLysGluArgGluLeuGluArgAsnArgGlnAlaAspIleGlu 1110
Qy 2173 GCCAAGCAGGCGCC---CAGGGTGTAACTTCTGCTTATATAGAAATGCCAAAGCC 2229
Db 1111 AlaLysArgAlaAlaArgGlnAspProLeuGluMetLeuIleGluGlyProLeuVal 1130
Qy 2230 -----CCA-----GAACCCAGACGACGTACACCTAGAA-----GACTGGCAA 2268
Db 1131 TrpGlyAsnProLeuTrpGluThrLysProGlnTyrlsSerGluGluIleGluTrp--- 1149
Qy 2269 GAGATAAAAAAGATAGACACGATCTCTGAGACTCCGGAAGGGACCTGCTATACCTCAGAT 2328
Db 1150 GlyIleSerArgGlyHisSerPhe-----LeuProSerGlyTrpGlnAlaThrGluGln 1167
Qy 2329 GGGAAGGAATCTCGCCCCCAAGAGAGGTGTAGATATGTCACACAGATACCTGCTA 2388
Db 1168 GlyLysIleProLeuProAlaAlaAsnGlnTrpLysLeuLeuLysThrHisGlnAsn 1187
Qy 2389 ACCCACCTA-----GGAACTAAACACCTGCAGCAGTTGGTCAGAACATCCCT--- 2436
Db 1188 PheHisLeuCysIleAspSerThrHisGlnMetAlaLysLeuLeuPheThrGlyProGly 1207
Qy 2437 ---TATCATGTTCTGAGGCTACCGAGTGGCTGACTCGGTGGTCAAAATGTGTGCC 2493
Db 1208 LeuPheLysThrLeuLys-----GlnIleValGlyAlaCysGluVal 1221
Qy 2494 TGCAGCTGGTTAATGTAAATCCT---TCCAGATGCTCCAGGGAACAGACTAAGGGA 2550
Db 1222 CysGlnIleAsnAsnProLeuProTyrlsHisGlnAlaProProGlyGluGlnArgThrGly 1241
Qy 2551 AGCCACCCCGGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGCTAAATACGGA 2610
Db 1242 HistySerGlyGluAspTrpGlnLeuAspPheThrHisMet---ProLysSerGlnGly 1260
Qy 2611 AACAAATACCTATTGGTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACT 2670
Db 1261 PheGlnTyrlsLeuLeuValTrpValAspIlePheThrGlyTyrlsAlaGluAlaPheProCys 1280
Qy 2671 AAGAAAGAGACTTCAACCGCTGCTTAAATAAATACTGGAAGAAATTTTCCAGATTT 2730
Db 1281 ArgThrGluLysAspGlnAspValIleLysAlaLeuIleHisGluIleProArgPhe 1300

Db	101	SerProAspArgThrTrpItyrThrValLeuAspLeuLysAspAlaPhePheCysLeuPro	120
Qy	685	TTACACCCCACTAGCCAAACCACTTTTTCCTTCCGAATGGAGAGATCCAGGTACCGGAGA	744
Db	121	LeuAlaProGlnSerGlnLeuPheAlaPheGluTrpArgAspProGluArgGlyIle	140
Qy	745	ACCGGCGAGCTCACCTGGACCCGACTGCCCCCAAGGGTTCAAGAACTCCCCGACCACTTT	804
Db	141	SerGlyGlnLeuThrTrpTrpArgLeuProGlnGlyPheLysAsnSerProThrLeuPhe	160
Qy	805	GACCAAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCACACCTCAGGTGACCCCTC	864
Db	161	AspGluAlaLeuHisArgAspLeuThrAspPheArgThrGlnHisProGluValThrLeu	180
Qy	865	CTCCAGTACGTGATGACTCTCTTCGCGGAGCCACCAACAGGACTGCTTAGAAGGC	924
Db	181	LeuGlnTyrrValAspAspLeuLeuAlaAlaProThrLysGluAlaCysIleArgGly	200
Qy	925	ACGAAGGCCTACTGCTGGAAATTGTCTGACTAGGCTACAGAGCCTCTGCTAAGAAGCC	984
Db	201	ThrLysHisLeuLeuArgGluLeuGlyAspLysGlyTyrrArgAlaSerAlaLysLeuAla	220
Qy	985	CAGATTTCAGGAGAGAGGTAAACATCTCTGGGTACAGTTTCGGGAGCGGCACCGATGG	1044
Db	221	GlnIleCysGlnThrTyrsValThrTyrrLeuGlyTyrrIleLeuSerGluGlyLysArgTrp	240
Qy	1045	CTGACGGAGCCGCGAAGAAACTGTAGTCCAGATACCGGCCCCCAACACACAGCCAAACAA	1104
Db	241	LeuThrProGlyArgIleGluThrValAlaHisIleProProGlnAsnProArgGlu	260
Qy	1105	ATGAGAGAGTTTTCGGGACAGCTGGATTTCAGACGTGTGGATCCCGGGTTCGCCACC	1164
Db	261	ValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGlu	280
Qy	1165	TTAGCAGCCCACTTACCCGCTAAACAAAGAAAAGGGAAATCTCTCGGCTCTCGAG	1224
Db	281	LeuAlaAlaProLeuTyrrAlaLeuThrLysGlnSerAlaProPheThrTrpGlnGluLys	300
Qy	1225	CACCAAGAGCATTTCGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCGGCCCTC	1284
Db	301	HisGlnSerAlaPheGluAlaLeuLysGluAlaLeuLeuSerAlaProAlaLeuGlyLeu	320
Qy	1285	CCTGACGCTAAACCCCTTTACCCTTTATGTGGATGAGCGTAAGGAGTAGCCCGGGA	1344
Db	321	ProAspThrSerLysProPheThrLeuPheIleAspGluLysGlnGlyIleAlaLysGly	340
Qy	1345	GTTTTAAACCAACCCCTAGACCATGGAGAAGACCTGTTCGCTTACTCTGTCAAGAAGCTC	1404
Db	341	ValLeuThrGlnLysLeuGlyProTrpLysArgProValAlaTyrrLeuSerLysLeu	360
Qy	1405	GATCCTGTAGCTGTGGCCCATATGCTCAAGGCTATCCAGCTGTGGCCATACG	1464
Db	361	AspProValAlaAlaGlyTrpProCysLeuArgIleMetAlaAlaThrAlaMetLeu	380
Qy	1465	GTCGAAGACCTGCACAAATTGACTTTGGGACAGAAATATACTGTAATAGCCCCCATGCA	1524
Db	381	ValLysAspSerAlaLysLeuThrLeuGlyGlnProLeuThrValIleThrProHisAla	400
Qy	1525	TTGAGAACATGCTTCGGAGCCGCCAGACCGATGATGATCAAC--GCCCGATGACC	1581
Db	401	LeuAlaAlaIleValArgGlnThrProAspArgTrpIleThrAsnAlaAlaArgLeuThr	420
Qy	1582	CACTATCAAGCTGCTCTC--ACAGAGAGGTACGTTGCTCCACAGCGGCTCTC	1638
Db	421	HisTyrrGlnAlaLeuLeuLeuAspThrAspArgIleGlnPheGlyProProValThrLeu	440
Qy	1639	AACCTCGCACTCTTCTGCTGAAGACACTGATGAACAGTGACTCATGATTGCCATCA	1698
Db	441	AsnProAlaThrLeuLeuProAlaProGluAspGlnSerAlaHisAspCysArgGln	460
Qy	1699	CTATTGATTGAGGAGACTGGGGTCCGAAGACCTT	1734
Db	461	ValLeuAlaGlnThrHisGlyThrArgGluAspLeu	472

RESULT 12

US-10-448-871A-26
; Sequence 26, Application US/10448871A
; Publication No. US20040093169A1
; GENERAL INFORMATION:
; APPLICANT: Mizusahi, Yoshiyuki
; APPLICANT: Sakaguchi, Kengo
; APPLICANT: Sugawara, Fumio
; TITLE OF INVENTION: A METHOD FOR DESIGNING A MOLECULAR STRUCTURE OF AN INHIBITOR FOR
; FILE REFERENCE: 03316/LH
; CURRENT APPLICATION NUMBER: US/10/448,871A
; CURRENT FILING DATE: 2003-05-29
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 26
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Baboon endogenous virus
US-10-448-871A-26

Alignment Scores:

Pred. No.:	2,33e-134	Length:	472
Score:	1768.50	Matches:	328
Percent Similarity:	81.3%	Conservative:	55
Best Local Similarity:	69.6%	Mismatches:	87
Query Match:	Indels:	1	
DB:	Gaps:	1	

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-448-871A-26 (1-472)

QY	325	GGTTGGCAAGCAAGTTCCTCCCAACAGTTATTCACAGTGAAGGCCAGTGCACACCAAGTG	384
DB	1	GlyArgAlaLysCysGlnAlaProIleIleAspLeuLysProThrAlaValProVal	20
QY	385	TCAGTCACACAGTACCCCTTCAGTAAGAAGCTCAAGAAGAAATTCGGCCCGCATGTCCAA	444
DB	21	SerIleLysGlnTyrProMetSerLeuGluAlaHisMetGlyIleArgGlnHisIleIle	40
QY	445	AGATTAAATCAACAGGGCATCTAGTTCCTGTCCAAATCTCCCTGGAATPACTCCCTCGTA	504
DB	41	LysPheLeuGluLeuGlyValLeuArgProCysArgSerProTyrAsnThrProLeuLeu	60
QY	505	CCGTTAGAAAGCTGGGACTAATGACTATTCGACAGTACAGGACTTCAGAGAGGTCAAT	564
DB	61	ProValLysLysProGlyThrGlnAspTyrArgProValGlnAspLeuArgGluIleAsn	80
QY	565	AAACGGGTGCAGGATATACACCAACAGTCCCGAACCTTATACTCTTGTGTGCTCTC	624
DB	81	LysArgThrValAspIleHisProThrValProAsnProTyrAsnLeuLeuSerThrLeu	100
QY	625	CCACCCCAACGGAGTGTGTATACAGTATGGAGCTTAAAGGATGCTCTTCTGCTCGTGA	684
DB	101	LysProAspTyrSerTyrThrValLeuAspLeuLysAspAlaPheCysLeuPro	120
QY	685	TTACACCCCACTAGCAACCACTTTTTCCTTCGATCGAGAGATCCAGGTACGGGAAGA	744
DB	121	LeuAlaProGlnSerGlnGluPheAlaPheGluTyrLysAspProGluArgGlyIle	140
QY	745	ACCGGGCAGCTCAGCTGGACCGCTGCCCAAGGGTTCAAGACTCCCGCCAGCATCTTT	804
DB	141	SerGlyGlnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrLeuPhe	160
QY	805	GACCAAGCCCTACACAGAGACTCGCCCAACTTCAGGATCCAAACCCCTCAGGTGACCCCTC	864
DB	161	AspGluAlaLeuHisArgAspLeuThrAspPheArgThrGlnHisProGluValThrLeu	180
QY	865	CTCCAGTACGTGGATGACTGCTGCTTCTGGGGAGCCACCAACAGGACTGCTAGAAGGC	924
DB	181	LeuGlnTyrValAspAspLeuLeuAlaProThrLysLysAlaCysThrGlnGly	200
QY	925	ACGAAGGCACACTGCTGGAATTGCTGACCTTAGGCTACAGAGCCTCTGCTAAGAAGGCC	984

DB	201	ThrArgHisLeuLeuGlnGluLeuGlyGlyLysGlyTyrArgAlaSerAlaLysLysAla	220
QY	985	CAGATTTGCAGGAGAGAGGTAAACATCTTGGGTACAGTTTGGGGGCGGCGGCGATGG	1044
DB	221	GlnIleCysGlnThrLysValThrTyrLeuGlyTyrIleLeuSerGlnGlyLysArgTrp	240
QY	1045	CTGACGGAGGCACGGAAGAAACCTGTAGTCAGATACCGGCCCAACACACAGCCAAACAA	1104
DB	241	LeuThrProGlyArgIleGluThrValAlaArgIleProProAsnProArgGlu	260
QY	1105	ATGAGAGAGTTTTTGGGGCAGAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTCGAC	1164
DB	261	ValArgGluPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaGlu	280
QY	1165	TTAGCAGCCCTCTATACCCGCTAACCAAGAAAGGGGAATTCCTCTGGGCTCTCGAG	1224
DB	281	LeuAlaAlaProLeuTyrAlaLeuThrLysGluSerThrProPheThrGlnThrGlu	300
QY	1225	CACCAGAAGGCATTTGATGCTATCAAAAAGCCCTGCTGAGCGCACCTGCTCTGCGCCTC	1284
DB	301	HisGlnLeuAlaPheGluAlaLeuLysLysAlaLeuLeuSerAlaProAlaLeuGlyLeu	320
QY	1285	CCTCAGCTAACTAAACCCCTTTTACCCCTTTATGTGTGAGCGGTAAAGGAGTAGCCGGGA	1344
DB	321	ProAspThrSerLysProPheThrLeuPheLeuAspGluArgGlnGlyIleAlaLysGly	340
QY	1345	GTTTTAAACCAACCCCTAGGACCAATGAGAGAGACCTGTCGCTACCTGTCAAGAAAGCTC	1404
DB	341	ValLeuThrGlnLysLeuGlyProTyrLysArgProValAlaTyrLeuSerLysLysLeu	360
QY	1405	GATCCTCTAGCCAGTGTGTCGCCCATATGCTGAAGCGTATCGCAGCTGTGGCCCATCTG	1464
DB	361	AspProValAlaAlaGlyTyrProCysLeuArgIleMetAlaAlaThrAlaMetLeu	380
QY	1465	GTCAGGACGCTGCAAAATTCAGCTTTGGGACAGAAATATACTGTAAATAGCCCCCATGCA	1524
DB	381	ValLysAspSerAlaLysLeuThrLeuGlyGlnProLeuThrValIleThrProHisThr	400
QY	1525	TTGAGAACATCTGTCGCCAGCCCCCAGACCGATGATGACCAACGGCGGATGACCCAC	1584
DB	401	LeuGluAlaIleValArgGlnProProAspArgTrpIleThrAsnAlaArgLeuThrHis	420
QY	1585	TATCAAGACCTGCTTCTC---ACAGAGAGGTACAGTTTCGCTCCACACGCGCTCTCAAC	1641
DB	421	TyrGlnAlaLeuLeuLeuAspThrAspArgValGlnPheGlyProProValThrLeuAsn	440
QY	1642	CCTGCCACTCTTCTGCTGCTGAAGAGACTGATGAACACAGTCACTCATGATTGCCATCACTA	1701
DB	441	ProAlaThrLeuLeuGluValProGluAsnGlnProSerProHisAspCysArgGlnVal	460
QY	1702	TTGATTGAGGAGACTGGGTCCCGCAAGGACCTT	1734
DB	461	LeuAlaGluThrHisGlyThrArgGluAspLeu	471

RESULT 13

US-10-450-763-33781
; Sequence 33781, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom

```

; SEQ ID NO 33781
; LENGTH: 1267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1066)..(1109)
; OTHER INFORMATION: CTF/NF-I proteins domain identified by eMATRIX, accession
; OTHER INFORMATION: number BL00349C, p-value=1.000e-09, raw score of 9.33
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (975)..(1120)
; OTHER INFORMATION: RNase H domain identified by Pfam, accession name rnaseH, E-
; OTHER INFORMATION: value=3.6e-34, Pfam score of 126.9
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1267)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-33781

Alignment Scores:
Pred. No.:      8.54e-107      Length:      1267
Score:          1431.50      Matches:      347
Percent Similarity: 55.6%      Conservative: 136
Best Local Similarity: 39.9%      Mismatches:  315
Query Match:    23.2%      Indels:       71
DB:              5          Gaps:         22

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-450-763-33781 (1-1267)

QY 52 GACTTGGGAGTGGACGGGTAAACCACTCGTTCTTCTGTCATACCTGAGTGCACGACCC 111
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 AspTrpGlyAlaLeuProPheSerHisAlaPheLeuIleMetProGluSerLeuThrPro 448
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 112 CTCTAGGTAGACATTATTACCAAGATGGAGCACAAATTTCTTTTGAACAGGNA 171
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 449 LeuLeuGluArgGluLeuValLysAlaGlyAlaIleHisLeuAsnIleGlyGlu 468
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 172 CCAGAAGTGTCTGCAATAACAAACCTATCATCTGTGTGACCTCCCAATTAGATGACGAA 231
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 469 -----GlyThrProIleCysArgLeuLeuPhe----- 477
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 232 TATCGACTATCTCTCCCTAGTAAAGCTGTATCAATATATACAAATTCGTGTTGGAACAG 291
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 -----GluGluGlyIleSer----- 482
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 292 TTTTCCCAAGCTCGGCAGAAACCGCAGGATGGTTTGGCAAGCAAGTTCCCCCAAA 351
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 483 ---ProGluValTrpAlaThrGluGlyGlnTyrGlyGlnAlaLysAsnAlaHisPheVal 501
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 352 GTTATTCAACTGAAGCCAGTGCACACACAGTGTCACTGACAGACAGTACCCTTGAGTAA 411
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 502 GlnValLysLeuLysAspSerThrSerPheProTyrGlnArgGlnTyrProLeuArgPro 521
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 412 GAAGCTCAAGAGAAATTCGGCCGCATGTCCAAAGATTAATCCAAAGGCGATCTTAGTT 471
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 522 GluAlaGlnArgLeuGlnLysIleValLysAspLeuLysAlaGlnGlyLeuValLys 541
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 472 CCTGTCCAATCTCCCTGGAATATCTCCCTGCTACCGGTAGAAAGCTGGGACTAATGAC 531
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 542 ProTyrSerProCysAsnThrProIleLeuGlyValGlnLysPro---LysArgGln 560
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 532 TATCGACCATGATACGAGGCTCAATAAAGCGGTGAGGATATACACCCCAACA 591
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 561 TrpArgLeuValGlnAspLeuArgIleIleAsnGluAlaValPheProLeuTyrProAla 580
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 592 GTCCGAAACCTTATAACCTCTGTGTGCTCTCCACCCCAACGAGCTGGTATACAGTA 651
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 581 IleProSerProTyrThrLeuLeuSerGlnIleProGluGluAlaGluTrpPheThrVal 600
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 652 TTGAGCTTAAGGATGCTCTCTCTGCTGAGATTACACCCACTAGCCCAACACTTTT 711
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 LeuAspLeuLysAspAlaPhePheCysIleProValHisProAspSerGlnPheLeuPhe 620
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 712 GCCTTCGAATGGAGAGATCCAGGTACGGAAGAAACCGGCGAGCTCACCTGGACCGCATG 771
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 621 AlaPheGlu-----AspPro---SerAsnProThrSerGlnLeuThrTrpThrValLeu 637
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 772 CCCCAAGGTTCAAGAACTCCCGACCATCTTTGACGAGCCCTACACAGACAGCTGGCC 831
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 638 ProGlnGlyPheArgAspSerProHisLeuPheGlyGlnAlaLeuAlaLeuAlaLeuSer 657
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 832 AACTTCAGGATCCAAACACCCCTCAGGTGACCTCTCCAGTACGTGGATGAGCTGCTCTG 891
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 GlnPhe-----SerTyrLeuAspThrProValLeuGlnCysMetAspAspLeuLeu 675
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 892 GCGGAGCCACCAACAGGACTCTTAGAAGCGACGAGGCACTACTCTCTGGAATTGTCT 951
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 676 AlaAlaArgSerGluThrLeuCysHisGlnAlaThrGlnAlaLeuLeuAsnPheLeuThr 695
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 952 GACCTAGGCTACAGAGCTCTGTGTAAGAGGCCAGATTTGACGAGAGAGGTAACTATAC 1011
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 696 ThrCysGlyTyrLysValSerLysProLysAlaGlnLeuCysSerGlnGlnValLysCys 715
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1012 TTGGGTTACAGTTTTCGCGGACGGGCGACGATGGCTGACGGAGGCGACGGAAGAACTGTA 1071
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 716 LeuGlyLeuLysLeuSerLysValThrArgAlaLeuSerGluGluArgIleGlnProIle 735
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1072 GTCCAGATACCGGCCCAACACACAGCCAAACAAATGACGAGAGTGTGTTGGGACAGCTGGA 1131
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 736 LeuAlaTyrProTyrProLysThrLeuLysGlnLeuArgGlyPheLeuGlyIleThrGly 755
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1132 TTTTGCAGACTGTGGATCCCGGGTTTGGCAGCTTACGAGCCCACTCTACCCGCTAAC 1191
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 756 PheCysArgIleTrpIleProArgTyrGlyLysIleAlaArgProLeuTyrThrLeuIle 775
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1192 AAGAAAAAGGGGAATTCCTCC-----TGGGCTCCTGAGCAGCAGCAGGCA 1236
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 776 LysGluThrGlnLysAlaAsnThrHisLeuValArgTrpThrProGluAlaGluAla 795
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1237 TTTGATGCTATCAAAAGGCGCTGCTGAGCGCACCTCTCTGGGCCCTCCCTGAGCTAACT 1296
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 796 PheHisAlaLeuLysLysAlaLeuMetGlnAlaProValLeuSerLeuLeu---ThrGly 814
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1297 AAACCTTTTACCCTTTATGTGGATGAGCGTAAG-----GGAGTA 1335
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 815 GlnAspPheSerTyrValThrLysAsnLysGlnThrLysLysLysLysLysLysLys 834
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1336 GCCCGGGGAGCTTTTAAACCCAAACCTAGGACCATGGAGAGACCTGTCGCTACTCTGCA 1395
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 835 AlaLeuArgValLeuAlaLeuVal**GlyThrSerLeuGlnProValAlaTyrLeuSer 854
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1396 AAGAAGCTCCTGATCCTGTAGCCAGTGGTTGGCCCATATGCTGAGAGGTATCGCAGCTGTG 1455
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 855 LysLysThrAspValValAlaLysGlyTrpProHisCysLeuTrpValMetAlaAlaIle 874
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1456 GCCATCTGTCTAAGGACGCTGACAAATTCATTTGGGACAGAAATATACTGTAAATAGCC 1515
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 875 AlaValLeuIleSerLysAlaValLysMetIleGln**ArgAspLeuThrValTrpThr 894
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1516 CCCCATGCTTGGAGAACATCGTTTCGGCAGCCCCCAGACCATGATGATGACACCGCCGC 1575
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 895 SerHisAspValAsnGlyIleLeuThrAlaLysGlyAspLeuTrpLeuSerAspAsnCys 914
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1576 ATGACCCACTATCAAAAGCCTCTTCTCACAGAGAGGTC---ACGTTTCGCTCCACCAAGCC 1632
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 915 LeuLeuLysCysGlnAlaLeuLeuLeuGluGlyProValLeuArgLeuCysThrCysAla 934
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1633 GCTCTCAACCTGCCACTCTTCTGCTGAGAGAGACTGATGAACAGTGAATCATGATGTC 1692
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 935 ThrLeuAsnProAlaThrPheLeuPro---AspAsnGluGluLysIleLysHisAsnCys 953
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1693 CATCACTATTGATTGAGGAGACTGGGGTCCGCAAGGACCTTACAGACTACACGCTGACT 1752
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 954 GlnGlnValIleSerGlnThrTyrAlaThrArgLysAspLeuLeuGluValProLeuThr 973
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


Db 151 GlnValLysLeuLysAspSerAlaSerPheProTyrGlnArgLysTyrProLeuArgPro 170
Qy 412 GAAGCTCAAGAGGAATTCGGCGCATGTGCAAGATTAAATCAACAGGCGATCCTAGTT 471
Db 171 GluAlaLeuGlnGly**GlnLysLeuValLysAspLeuLysAlaGlnGlyLeuValLys 190
Qy 472 CCTGTCCAATCTCCCTGGAATACCTCCCTGTGTACCGGTAGAAAGCTGGGACTAATGAC 531
Db 191 ProCysSerSerProCysAsnThrProLeuGlnGlyValArgLysPro---AsnGlyGln 209
Qy 532 TATCCAGCATACAGGACTTGAGAGAGGTCAATAAAGCGGTGCAGGATATACACCAACA 591
Db 210 TrpArgLeuValGlnAspLeuArgIleIleAsnGluAlaValPheProLeuTyrProAla 229
Qy 592 GTCCGAAACCTTATAACCTCTGTGTGTCTCCACCCCAACGAGGAGTGTGTATACAGTA 651
Db 230 ValSerSerProTyrThrLeuLeuSerLeuIleProGluGluAlaGluTrpPheThrVal 249
Qy 652 TTGAGCTTAAAGGATGCTCTCTGCTGTGAGATTACACCCACTAGCCACCACTTTT 711
Db 250 LeuAspLeuLysAspAlaPheCysIleProValArgProAspSerGlnPheLeuPhe 269
Qy 712 GCCTTCGAATGAGAGATCCAGGTACGGGAAGAACCGGCGAGCTCACCTGACCGACTG 771
Db 270 AlaPheGlu-----AspPro---LeuAsnProThrSerGlnLeuThrTrpThrValLeu 286
Qy 772 CCCAAGGTTCAAGACTCCCGACCACTCTTGACGAAGCCCTACACAGAGACTGGCC 831
Db 287 ProGlnGlyPheArgAspSerProHisLeuPheGlyGlnAlaLeuAlaGlnAspLeuSer 306
Qy 832 AACTTCAGGATCCAAACCCCTCAGTGACCTCTCCAGTACGTGGAGTACCTGCTCTG 891
Db 307 GlnPhe-----SerTyrLeuAspThrLeuValLeuGlnTyrValAspAspLeuLeu 324
Qy 892 GCGGGAGCCACCAACAGGACTGTAGAGCGCAGGAGGCACTACTGCTGGAATGCT 951
Db 325 ValAlaArgSerGluThrLeuCysHisGlnAlaThrGlnGluLeuLeuThrPheLeuThr 344
Qy 952 GACTAGGCTACAGAGCTCTGTGTAAGAGCCAGATTTGACGAGAGAGGTAAACATAC 1011
Db 345 ThrCysGlyTyrLysValSerLysProLysAlaArgLeuLysSerGlnGluLeuArgTyr 364
Qy 1012 TTGGGTACAGTTTGGGCGGCGGAGGCTGTGAGGCTGAGGAGGCGACGAGAAACGTGA 1071
Db 365 LeuGlyLeuLysLeuSerLysGlyThrArgAlaLeuSerGlnGluArgIleGlnProIle 384
Qy 1072 GTCCAGATACCGGCCCAACACACAGCAAAATATGAGAGATTTTGGGCGACAGCTGA 1131
Db 385 LeuAlaTyrProHisProLysThrLeuLysGlnLeuArgGlyPheLeuGlyIleThrGly 404
Qy 1132 TTTTGCAGACTGTGGATCCCGGGTGTTCGACCTTAGCAGCCCACTCTACCCGCTAAC 1191
Db 405 PheCysArgLysGlnIleProArgTyrThrProIleAlaArgProLeuTyrThrLeuIle 424
Qy 1192 AAGAAAAGGGGAATTCCTCC-----TGGGCTCTGAGCACCAGAGGCA 1236
Db 425 ArgGluThrGlnLysAlaAsnThrTyrLeuValArgTrpThrProThr---GluValAla 443
Qy 1237 TTTGATCTATCAAAAGGCGCTCTGAGCGCAGCTCTGCGCCCTCCCTGAGCTAACT 1296
Db 444 PheGlnAlaLeuLysLysAlaLeuThrGlnAlaProValPheSerLeuPro---ThrGly 462
Qy 1297 AAACCTTTTACCTTTATGTGATGAGCGTAAGGAGTAGCCGGGAGTTTAAACCAA 1356
Db 463 GlnAspPheSerLeuTyrAlaThrGluLysThrGlyIleAlaLeuGlyValLeuThrGln 482
Qy 1357 ACCTAGGACCATGAGAGACCTGTGCGCTTACCTGTCAAGAGAGCTCGATCTGTAGCC 1416
Db 483 ValSerGlyWetSerLeuGlnProValValTyrLeuSerLysGluIleAspValValAla 502
Qy 1417 AGTGGTTGGCCCATATCCCTGAAGGCTATCGAGCTGTGGCCATCTGTGTCAGGAGCT 1476
Db 503 LysGlyTrpProHisCysLeuTrpValMetAlaValAlaValLeuValSerGluAla 522

Qy 1477 GACAAATTGACTTTTGGGACAGATATAACTGTAATAGCCCCCATGCAATTGGAGAACATC 1536
Db 523 VallylIleIleGlnGlyArgAspLeuThrValTrpThrSerHisAspValAsnGlyIle 542
Qy 1537 GTTCGGCAGCCCCCAGACCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1596
Db 543 LeuThrAlaLysGlyAspLeuTrpLeuSerAspAsnHisLeuLeuAsnTyrGlnAlaLeu 562
Qy 1597 CTTCTCACAGAGGGTCT---ACGTTGCTCCACAGCGCTCTCAACCTCGCCACTCTT 1653
Db 563 LeuLeuGluGluProValLeuArgLeuArgThrCysAlaThrLeuLysProAlaThrPhe 582
Qy 1654 CTGCTGAAGAGACTGATGAACACGAGTACTCATGATGATGATGATGATGATGATGATGATG 1713
Db 583 LeuPro---AspAsnGluGluLysIleGluHisAsnGlnGlnValIleAlaGlnThr 601
Qy 1714 ACTGGGTCGCGCAGGACCTTACAGACATACCGCTGACTGAGTGTGAGTGTGACTACCTGGTTC 1773
Db 602 TyrAlaAlaArgGlyAspLeuLeuGluValProLeuThrAspProAspLeuAsnLeuTyr 621
Qy 1774 ACTCACGAAGCAGCTATGTGCGAAGGTAAAGAGTGGCTGGCGCGCGTGTGTGAC 1833
Db 622 ThrAspGlySerSerLeuAlaGluLysGlyLeuArgLysAlaGlyTyrAlaValIleSer 641
Qy 1834 GGGACCCGCGCAGCTCTGGCGCAGCAGCTGCGGGAAGAACTTCAGCACAAAAGCTGAG 1893
Db 642 AspAsnGlyIleLeuGluSerAsnArgLeuThrProGlyThrSerAlaHisLeuAlaGlu 661
Qy 1894 CTCATGCGCTCAGCAAGCTTTGCGCTGCGCGCAGGGAATCCATAACATTTATACG 1953
Db 662 LeuIleAlaLeuThrTrpAlaLeuGluLeuGlyGlyArgValAsnIleTyrSer 681
Qy 1954 GACAGAGGTATGCTTTGCGACTGTCACATGCGGCGCATCTATAACAAGGGGG 2013
Db 682 AspSerLysTyrAlaTyrLeuValLeuHisAlaHisAlaIleTrpArgGluArgGlu 701
Qy 2014 TTGCTTACCTCAGCAGGAGGAATAAAGAACAAAGAGGAATTTAAAGCTTATTGAA 2073
Db 702 PheLeuThrSerGluGlyThrProIleAsnHisGlnGluAlaIleArgArgLeuLeu 721
Qy 2074 GCGGTACATTTACAAAAGCTTAGCTATTATACATCTGCTGCTGAGCATCAGAAAGCTAAA 2133
Db 722 AlaValGlnLysProLysGluValAlaValLeuHisCysGlnGlyHisGlnGluGlu 741
Qy 2134 GATCTCATATCCAGAGGAACACAGATGCTGCGGCTTCCAGAGCGGAGCCAG 2190
Db 742 GluArgGluIleGluGlyAsnArgGlnAlaAspIleGluAlaLysLysAlaAlaArg 760

RESULT 15

US-10-114-104-89
; Sequence 89, Application US/10114104
; Publication No. US20030198647A1

GENERAL INFORMATION:

APPLICANT: PERRON, HERVE
BESEME, FREDERIC
BEDIN, FREDERIC
PARANHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUBE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA

```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,104
; FILING DATE: 03-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,847
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 768 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030198647A1 Relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 89:
US-10-114-104-89

Alignment Scores:
Pred. No.: 138e-105 Length: 768
Score: 1415.50 Matches: 318
Percent Similarity: 58.9% Conservative: 117
Best Local Similarity: 43.0% Mismatches: 255
Query Match: 22.9% Indels: 49
DB: 4 Gaps: 14

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-114-104-89 (1-768)
QY 4 GGTGCCACAGGGCAACACAGTATCCATGGACTACC-----CGAAGAACAGTTT 51
Db 61 GlyValLeuGlyGln-----ProValThrThrTyrrPheSerGlnProLeuSerCys 77
QY 52 GACTTGGAGTGGGACGGGTACCCACCTGTTCTGTGTCATACCTGAGTGGCCAGCACCC 111
Db 78 AspTrpGlyThrLeuLeuPheSerHisAlaPheLeuIleMetProGluSerProThrPro 97
QY 112 CTCTTAGGTAGAGCTTATTGACCAAGATGGGAGCACAAATTTCTTTTGAACAGGGAAA 171
Db 98 LeuLeuGlyArgaspIleLeuAlaLysAlaGlyAlaIleIleHisLeuAsnIleGlyLys 117
QY 172 CCAGAAGTGTCTGCAAAATAACAAACCTATCACTGTGTGTTGACCCCTCAATTAGATGACGA 231
Db 118 -----GlyIleProIleCys----- 122
QY 232 TATGACTATATCTCTCCCTAGTAAGCCTGATCAAAATATACAAATTTCTGGTTGGAAACAG 291
Db 123 -----CysProLeuLeuGluGlyIleAsn----- 131
QY 292 TTTCCCAAGCTGGGCGAACCAGCAGGGATGGGTTTGGCAAGCAAGTTCCCCCAACA 351
Db 132 ---ProGluValTrpAlaIleGluGlnTyrrGlyGlnAlaLysAsnAlaArgProVal 150
QY 352 GTTATTAACTGAAGCCAGTGCCACACAGTGTGTCAGTCAGACAGTACCCCTTGAGTAAA 411
Db 151 GlnValLysLeuLysAspSerAlaSerPheProTyrrGlnArgLysTyrrProLeuArgPro 170
QY 412 GAAGCTCAAGAAAGTAATTCGGCCGATGTCGAAAGATTATCCACAGGGCATCTCTAGTT 471
Db 171 GluAlaLeuGlnGly***GlnLysIleValLysAspLeuLysAlaGlnGlyLeuValLys 190
QY 472 CCTGTCCAATCTCCCTGGGAATACTCCCTGCTACCGGTTAGAAAGCCTGGGACTAATGAC 531
```

```
Db 191 ProCysSerSerProCysAsnThrProIleLeuGlyValArgLysPro---AsnGlyGln 209
QY 532 TATCGACCACTACAGGACTTGCAGAGGTCAATAAACGGGTGCAGGATATACACCCCAACA 591
Db 210 TrpArgLeuValGlnAspLeuArgIleIleAsnGluAlaValPheProLeuTyrrProAla 229
QY 592 GTCCCGAAACCTTATAACCTCTTGTGTGCTCTCCACCCCAACCGAGCTGTATACAGTA 651
Db 230 ValSerSerProTyrrThrLeuLeuSerLeuIleProGluGluAlaGluTrpPheThrVal 249
QY 652 TTGGACTTAAAGGATGCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTT 711
Db 250 LeuAspLeuLysAspAlaPhePheCysIleProValArgProAspSerGlnPheLeuPhe 269
QY 712 GCCTTCGAATGGAGATCCAGGTACCGGAAGACCGGGAGAGCTCACCTGGACCGGACTG 771
Db 270 AlaPheGlu-----AspPro---LeuAsnProThrSerGlnLeuThrTrpThrValLeu 286
QY 772 CCCCAAGGTTCAAGAACTCCCCGACCATCTTTTACGAAGCCCTACACAGAGACTGGCC 831
Db 287 ProGlnGlyPheArgAspSerProHisLeuPheGlyGlnAlaLeuAlaGlnAspLeuSer 306
QY 832 AACTTCAGGATCCCAACACCTCAGGTGACCTCTCAGTACGTGAGTACCTGCTTCTG 891
Db 307 GlnPhe-----SerTyrrLeuAspThrLeuValLeuGlnTyrrValAspLeuLeuLeu 324
QY 892 CGGGAGCCCAACAAACAGGACTGCTTAGAAGGCACGAAGGCACCTACTGCTGGAATTGCT 951
Db 325 ValAlaArgSerGluThrLeuCysHisGlnAlaThrGlnGluLeuLeuThrPheLeuThr 344
QY 952 GACCTAGCTACAGAGCTCTGCTAAGAAGCCCAAGATTTTCAGAGAGAGAGGTAAACATAC 1011
Db 345 ThrCysGlyTyrrLysValSerLysProLysAlaArgLeuCysSerGlnGluIleArgTyrr 364
QY 1012 TTGGGGTACAGTTTTCGGGACGGGACGATGGCTGACGGAGGCACGGAAGAAACTGTA 1071
Db 365 LeuGlyLeuLysLeuSerLysGlyThrArgAlaLeuSerGluGluArgIleGlnProIle 384
QY 1072 GTCAGATACCGGCCCCCAACCAACCAAGCAATAGAGAGTGTTTTGGGACAGCTGGA 1131
Db 385 LeuAlaTyrrProHisProLysThrLeuLysGlnLeuArgGlyPheLeuGlyIleThrGly 404
QY 1132 TTTTGCAGACTGTGGATCCCGGGTTTTCGACCTTAGCAGCCCACTCTACCCGCTAAC 1191
Db 405 PheCysArgLysGlnIleProArgTyrrThrProIleAlaArgProLeuTyrrThrLeuIle 424
QY 1192 AAAGAAAAGGGGAATTTCTCC-----TGGGCTCTCTGACGACACAGAGGCA 1236
Db 425 ArgGluThrGlnLysAlaAsnThrTyrrLeuValArgTrpThrProThr---GluValAla 443
QY 1237 TTTGATGCTATCAAAAAGGCCCTGCTGAGCGCAGCTGCTCTGGCCCTCCCTGACGTA 1296
Db 444 PheGlnAlaLeuLysLysAlaLeuThrGlnAlaProValPheSerLeuPro---ThrGly 462
QY 1297 AAACCCCTTACCTTTATGTGATGAGGTAAAGGAGTAGCCCGGGAGTTTAAACCAA 1356
Db 463 GlnAspPheSerLeuTyrrAlaThrGlnLysThrGlyIleAlaLeuGlyValLeuThrGln 482
QY 1357 ACCCTAGACCATCGAAGAGACCTGTCGCTTACCTGTCTCAAAAGAGCTCGATCTGTAG 1416
Db 483 ValSerGlyMetSerLeuGlnProValValTyrrLeuSerLysGluIleAspValValAla 502
QY 1417 AGTGGTGGCCCATATGCTCGAAGCTATCCAGCTGTGGCCATCTGGTCAAGGACCT 1476
Db 503 LysGlyTrpProHisCysLysLeuTrpValMetAlaAlaValAlaValLeuValSerGluAla 522
QY 1477 GACAAATTTGACTTTGGGACAGATATAACTGTAAATAGCCCCCATGCTTGGAGAACATC 1536
Db 523 ValLysIleIleGlnGlyArgAspLeuThrValTrpThrSerHisAspValAsnGlyIle 542
QY 1537 GTTCGGCAGCCCGCAGATGATGATGACCAACCGCCGATGACCCACTATCAAGCCTG 1596
Db 542
```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:19:01 ; Search time 3.01216 Seconds
(without alignments)
2992.933 Million cell updates/sec

Title: US-10-723-552-3_COPY_2307_5741

Perfect score: 6183

Sequence: 1 ATGGGTCCACAGGCAACA.....CTGTCAATAACCTCTCAGAC 3435

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 194028

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10723552/runat 14022006 125151 13368/app_query.fasta_1
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN.1 1 17 @runat 14022006 125151 13368 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:
5: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pcp:
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp:
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp:
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1660.5	26.9	1189	6	US-10-821-234-1209
2	1612	26.1	1150	6	US-10-821-234-1083
3	431	7.0	912	7	US-11-042-988-12
4	301	4.9	562	6	US-10-507-928-10
5	301	4.9	562	6	US-10-507-928-12
6	301	4.9	562	7	US-11-029-465-10
7	287.5	4.6	561	7	US-11-029-465-12
8	217.5	3.5	210	7	US-11-022-562-221
9	210.5	3.4	422	7	US-11-230-251-26

10	174.5	2.8	188	7	US-11-234-786-592
11	146.5	2.4	14130	7	US-11-175-689-9
12	143.5	2.3	16990	7	US-11-175-689-7
13	126.5	2.0	5712	7	US-11-143-980-47
14	126	2.0	745	6	US-10-995-561-659
15	126	2.0	745	6	US-11-109-156-14
16	126	2.0	745	6	US-11-222-158-4
17	124.5	2.0	937	7	US-11-017-550-66
18	121.5	2.0	2410	7	US-11-175-689-8
19	121.5	2.0	7968	7	US-11-186-731-5
20	121	2.0	900	7	US-11-182-016-37
21	121	2.0	2591	6	US-10-453-372-718
22	121	2.0	2602	6	US-10-453-372-716
23	121	2.0	2617	6	US-10-453-372-666
24	121	2.0	2617	6	US-10-453-372-732
25	121	2.0	2617	6	US-10-453-372-734
26	121	2.0	2617	6	US-10-453-372-736
27	121	2.0	2617	6	US-10-453-372-738
28	121	2.0	2617	6	US-10-453-372-740
29	121	2.0	2617	6	US-10-453-372-742
30	121	2.0	2617	6	US-10-453-372-744
31	121	2.0	2617	6	US-10-453-372-748
32	121	2.0	2617	6	US-10-453-372-750
33	120.5	1.9	2766	6	US-10-877-346-62
34	120	1.9	1717	7	US-11-182-016-20
35	119.5	1.9	216	6	US-10-821-234-984
36	119.5	1.9	6893	7	US-11-205-109-14
37	119	1.9	2617	6	US-10-453-372-746
38	118.5	1.9	1368	7	US-11-043-693-34
39	118	1.9	1709	6	US-10-995-561-973
40	117.5	1.9	919	6	US-10-858-730-206
41	117	1.9	1742	7	US-11-182-016-23
42	116.5	1.9	1439	7	US-11-124-368A-291
43	116.5	1.9	1649	6	US-10-985-561-974
44	114.5	1.9	1560	7	US-11-059-962-1
45	114.5	1.9	2768	6	US-10-510-101-72

ALIGNMENTS

RESULT 1
US-10-821-234-1209
; Sequence 1209, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1209
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1189)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-821-234-1209

Alignment Scores:
Pred. No.: 3.02e-122 Length: 1189
Score: 1660.50 Matches: 431
Percent Similarity: 51.3% Conservative: 184
Best Local Similarity: 35.9% Mismatches: 445
Query Match: 26.9% Indels: 139

[illegible]

Db	372	MetMetThrTyrLeuGlyIleGlnLeuSerProGlyAlaGlnAlaMetThrProAlaArg	391
Qy	1060	AAGAAAACTCTAGTCAGATACCGGCCCAACACACAGCCAAACAAATAGAGAGAGTTTGTG	1119
Db	392	AlaAlaLeuIleGlnAsnLeuProProSerSerLeuSerGluIleLeuSerPhePro	411
Qy	1120	GGGACAGCTGGATTTTCAGACTGTGGATCCCGGGGTTTCGCAGCTTTAGCAGCCCACTC	1179
Db	412	GlyLeuAlaGlyPhePheArgIleTyrPheSerAsnPheAlaLeuLeuAlaHisProLeu	431
Qy	1180	TACCGGCTAACCAAGAAAAAGGGGAATTCCTCTGGGCTCCTGAGCAGCAG-----AAG	1233
Db	432	TyrGluValAlaIleGlyProProAsnGluProLeuAsnProSerHisAsnIleLeuPro	451
Qy	1234	GCATTTGATGCTATCAAAAGGCGCTGTGAGCGCACCTGCTCTGGCCCTCCCTGACGTA	1293
Db	452	SerPheHisLeuSerGlnThrAlaLeuValThrAlaSerAlaProSerLeuProAspIle	471
Qy	1294	ACTAAACCCCTTTACCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGGGGAGTTTAAAC	1353
Db	472	SerGlnProPheThrLeuTyrThrAlaGluSer**GlyIleAlaLeuGlyValLeuGly	491
Qy	1354	CAAAACCCCTAGGA-----CCATGGAGAAGACCTGTGCGCTACTCTCAAGAAGCTCGAT	1407
Db	492	GlnGlnIleGlyAsnProProSerPheAlaProValAlaTyrLeuSerLeuGlnLeuAsp	511
Qy	1408	CCTGTAGCAGCTGTTGGCCCATATGCTGAAGGCTATCGAGCTCTGGCCATACATGGTC	1467
Db	512	AsnThrValIleGlyIleProAlaCysPheLeuAlaLeuGluValAlaAlaSerLeuAla	531
Qy	1468	AAGAGCGCTGACAAATTTGACTTTGGGACAGAAATATAACTGTATAGCCCCCATGATTG	1527
Db	532	LeuGluSerArgIleLeuThrPheSerGlnAsnThrThrValHisSerHisAsnLeu	551
Qy	1528	GAGAACATCTGTCGGCAG-----CCCCCAGACGATGGATGAGCAAC	1569
Db	552	GlnAspLeuLeuSerSerGlnAlaValSerSerLeuProProSerArg-----	567
Qy	1570	GCCGCGATGACCCACTATCAAGCGCTCTTCACAGAGGGGTACGTTGCGTCCACCA	1629
Db	568	IleGlnLeuIleHis---AlaLeuPheIleLeuAsnProIlePheSerLeuThrArgSer	586
Qy	1630	GCGCTCTCAACCGCTGCACCTCTTCTGCTGGAAGAGACTGATGAACACGATCATGAT	1689
Db	587	AlaSerLeuAsnProAlaSerLeuLeuProValSerSerSerLeuPro---ThrHisSer	605
Qy	1690	TGCCATCACTATTGATTAGGAGACTGGGTCGCCAAGACCTTACAGACATACCGCTG	1749
Db	606	SerThrAspIleLeuAspHisValGlnProHisPheProAsnThrSerGluProLeu	625
Qy	1750	ACTGGAGAGTCTAACCTGGTTCACTGACGGAGCAGCTATGTGTGGTGAAGGTAAAGG	1809
Db	626	ThrAsnProAsnAspGlnLeuPheIleAspGlySerSerPheGlnAla-----His	642
Qy	1810	ATGCGTGGGCGGGTGTGGACGGACCGCACGATCTGGCCAGCAGCTTCGCGGAA	1869
Db	643	GlnLeuProProArgLeuLeuAspMetGlnLeuPheProLeuThrIleLeuLeuPro	662
Qy	1870	GGAACTTCAGCAAAAGGCTGAGCTCATGGCCCTCACGAAGCTTTTGGCGTCCGCGAA	1929
Db	663	GlySerSerSerGlnIleAlaGluLeuIleAlaLeuThrArgAlaLeuAsnLeuSerLys	682
Qy	1930	GGGAAATCCATAACATTTATACGGACAGCAGGTATGCCCTTCCGACTGCACAGTACAT	1989
Db	683	GlyLysArgValAsnIleTyrThrAspSerLysIleTyrAlaTyrHisIleProArgSerHis	702
Qy	1990	GGGGCCATCTATAACAAAGGGGTGTGCTTACCTCAGCAGGAGGAGGAAATAAGACAA	2049
Db	703	AlaAlaIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle	722
Qy	2050	GAGGAAATCTTAAGCCCTATTAGAAGCGGTACATTTTACCAAAAGGCTAGCTATTATAC	2109

Db 723 HisLeuLeuTyr***LeuLeuGlnAlaThrHisLeuProAlaLysAlaGlyValIleHis 742
Qy 2110 TGCTCTGACATCAAGAGCTAAAGATCTCATATCCAGAGAAACACGATCGCTGACCGG 2169
Db 743 Cys**GlyHisArgThrGlySerAspGluIleSerLysGlyAsnArgLysThrAspGlu 762
Qy 2170 GTTCCAAAGCAGCCAGCCAGCGGTGTAACTCTCTGCTCTAATAAGTAAGTCCCAAGGCC 2229
Db 763 AlaAlaLysGlnAspSerLeuSerProLeuProAlaProIleLeuLeuValThrProAla 782
Qy 2230 CCAGAACCCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2289
Db 783 ValProProArgTyrProProThr-----GluLysSerSerLeuLeuGln 797
Qy 2290 TTCTCTGAGACTCCGGAAGGACCTGCTATACCTCAGATGGAAGAAATCTCGCCCCAC 2349
Db 798 GlnGlyAlaSerLeuGlnGlyAspTrpIleIleLysAsnGlnLysProValLeuProGln 817
Qy 2350 AAGAAGGTTAGAAATGTCCAACAGATACATCGTCTTAACCCACCTAGGAACTAAACAC 2409
Db 818 GluGlnSerLysGluIleLeuThrProLeuHisGlnProPheHisIleSerAlaCysPro 837
Qy 2410 CTGCAGCAGTGGTCAG-----ACATCCCTTATCATGTCTGTGAGGCTACGAGGA 2460
Db 838 LeuTyrLeuLeuLeuArgProCysPheSerSerProHisLeuPheIleSerLeu----- 855
Qy 2461 GTGCTGACTCGTGGTCAAAACATGTGTGCCCTGCCAGCTGGTGAATGCTTAATCCTTC 2520
Db 856 -----LysAspIleThrSerAsnCybArgIleCysSerValThrSerSerGlnGlyAla 873
Qy 2521 AGAATGCTT-----CCAGGAAGAGACTAAGGGGAAGCCACCCAGCGGCTCCTACGG 2571
Db 874 LeuCysProLeuLeuLeuLeuTyrGlnLeuArgGlyThrLeuProGlyGluHisTrp 893
Qy 2572 GAATGAGCTTCACTGAGTAAAGCGGCTAAATACGGAACAAATACCTATTGTTGTTTT 2631
Db 894 GlnValAsnPheThrHisMetProProValLys---LysSerLysTyrLeuLeuThrLeu 912
Qy 2632 GTAGACACCTTTTCAGGATGGGTAGAGCTTATCTACTAAGAAAGAGACTTCAACCGTG 2691
Db 913 ValAspThrPheSerGly***ValGluAlaPheProThrProSerGlnLysAlaAlaGlu 932
Qy 2692 GTGCTAAAAAATACCTGGAAGAAATTTTCCAAAGATTTGGAATPACCTAAGGTAATAGG 2751
Db 933 ValSerGlnIleLeuValThrGluIleIleProArgPheGlyLeuProGlySerIleGln 952
Qy 2752 TCAGACATGGTCCAGCTTTTGTGCCAGGTAGTCAAGGACGAGGACGAGATATCGGG 2811
Db 953 SerAspAsnSerProSerPheIleSerGlnIleThrGlnGlnValSerGlnSerLeuGly 972
Qy 2812 ATTGATTGGAACTGCATTGTGCATACAGACCCCAAGAGCTCAGACAGGTAGAGAGCATG 2871
Db 973 IleGlnTrpArgLeuHisIleProCysTrpProGlnThrSerGlyLysValGluArgAla 992
Qy 2872 AATAGAACCATTAAGAGACCTTCTAATAATGTACCGCGGAGCTGCGCTTAATGATGTA 2931
Db 993 AsnGlyIleLeuLysAlaGlnLeuThrLysLeuThrLeuGluVal---GlnLysProTrp 1011
Qy 2932 ATAGCTCTCTGCTGCTTTTGTGCTTTTGTAGGTTAGGAACCCCT---GGACAGTTTGGG 2988
Db 1012 ThrSerLeuLeuProIleAlaLeuGluSerIleArgAlaSerProLysAlaProSerPhe 1031
Qy 2989 CTGACCCCTTATGAATTAATCTACGCGGGGACCC-----CCCCCATTTG 3030
Db 1032 LeuSerProPheGluLeuIleTyrGlyArgProPheLeuLeuGlnAsnArgProPro--- 1050
Qy 3031 GTAGAAATGCTCTGTACATAGTGTGACGTGCTGCTTTTCCAGCCTTTGTTCTCTAGG 3090
Db 1051 -----SerAsnSerGlnLeuGlyGluTyrLeuProThrValSerLeuMetSerTyr 1067
Qy 3091 CTCNAGGCATTTGAGTGGTGAGACACGAGCGTGAGGCACTCCGGGAGCGCTACTCA 3150
Db 1068 Leu-----LeuCysGlnGlnAlaAspGlnAlaLeuProLysProHisGlu 1082

Qy 3151 GGAGGAGGAGACTTGCAG-----ATCCACATCGTTTCCAAGTGGAGAT 3195
Db 1083 GlyValSerAsnProLys***ThrCysSerProIleProLys-----Asp 1097
Qy 3196 TCA-----GTCCTACGTTAGACGACCGCTCAGAGAAACCTCGAGACTCGGTGGGAG 3246
Db 1098 SerLeuSerArgValThrLeuGlnAsnHisArgGlyLeuAspLeuLeuThrAlaGluLys 1117
Qy 3247 GGCCCTTATCTCGTACTTTTG-----ACCACACCAACGGCTGTG 3285
Db 1118 GlyGlyLeuCysIlePheLeuGluGluCysCysPheTyrThrAsnGlnSerGlyLeu 1137
Qy 3286 -----AAAGTCGAAGGAATCTCCACCTGGATCCAT 3315
Db 1138 ValGlnAspAlaAlaGlyArgIleAsnGlnLysAlaSerGlyArgValGlnTrpLeu--- 1156
Qy 3316 GCATCCACCGTTAAACCGCGCCACCTCCCGATTGGGGTGGAAAGCGGAAAG----- 3369
Db 1157 -----ThrProValIleProAspLeuTrpGluAlaGluAlaGlyGly 1170
Qy 3370 -----ACTGMAAATCCCTTAAGCTTCGCTCCAT 3399
Db 1171 SerArgGlyGlnGluIleGluThrIleLeuAlaAsnThrValLysProArgLeuTyr 1189

RESULT 2

US-10-821-234-1083
; Sequence 1083, Application US/10821234
; Publication No. US20050255114A1

; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom

; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234

; CURRENT FILING DATE: 2004-04-07

; PRIOR APPLICATION NUMBER: US 60/462,047

; PRIOR FILING DATE: 2003-04-07

; NUMBER OF SEQ ID NOS: 1704

; SOFTWARE: pt_seq_genes Version 1.0

; SEQ ID NO 1083

; LENGTH: 1150

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)..(1150)

; OTHER INFORMATION: Xaa = any amino acid or nothing

US-10-821-234-1083

Alignment Scores:

Pred. No.: 1,86e-118 Length: 1150

Score: 1612.00 Matches: 421

Percent Similarity: 52.3% Conservative: 178

Best Local Similarity: 36.8% Mismatches: 444

Query Match: 26.1% Indels: 102

DB: 6 Gaps: 35

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-821-234-1083 (1-1150)

Qy 82 TTCTCTGTCATACCTGAGTCCCGACACCCCTCTTAGGTAGAGACTTATTGACCAAGATG 141

Db 59 PheLeuCysProGluCysProThrProLeuLeuGlyArgAspLeuLeuSerLysMet 78

Qy 142 GGACGACAAATTTTGTGAACAAAGGAAACCAAGAGTGTCTGCAAT---AACAAACCT 198

Db 79 GlyAlaThrIleSerLeuGluGluAspArgLeuGlnValGluAlaGluProGluGlnGly 98

Qy 199 ATCACTGTGTTCACCTCCCAATTAGATGACGAATATGACTATCTCTCCCTCTAGTAAG 258

Db 99 IleHisLeuLeuAlaLeuAsnGlyGlnGluLeuGlu----- 111


```
Db 810 -----*TrpArgAspValLeuTrpAsnSerValHisLysLysPheThrArg 825
Qy 2455 CCAGAGTGGCTGACTCGGTGGTCAAAACATTGTGCTCCAGCTGCTTAATGCTAAT 2514
Db 826 LysLysLeuArgLysSerValLysGlnValThrPheValSerGluLeuCysSerArgAsn 845
Qy 2515 CCTTCCAGA-----ATGCCTCCAGGGAAG-----AGACTAAGGGGAAGC 2553
Db 846 AsnProHisThrHisProIleProProSer***LeuLysLeuValGlnHisArgGlyThr 865
Qy 2554 CACCCAGCGCTCACTCGGAGTGGACTTCACTGAGGTAAAGCGGCTTAATACGGAAC 2613
Db 866 TyrLeuGlyGluAspTrpGlnValValLeuThrGlnMetThrPro---AsnLeuGlyTyr 884
Qy 2614 AAATACCTATTGGTTTTGTAGACACCTTTTTCAGATGGGTAGAGGCTTATCCTACTAAG 2673
Db 885 LysTyrLeuLeuValPheValAspThrPheThrGlyArgValLysGlyPheProThrCys 904
Qy 2674 AAAGAGACTTCAACCGTGGTGGCTAAATAAATACTGGAAGAAATTTTCCAGATTTGGA 2733
Db 905 ThrGluLysAlaValGluValCysLysProSerLeuLysGluValIleSerGlnPheGly 924
Qy 2734 ATACCTAAGGTAAAGGTGACAACTGTCAGCTTTTGTGCTCCAGGTAAAGTCAGGGA 2793
Db 925 LeuProLysSerProGlnSerGlyAsnArgLeuSerPheMetGlyLysIleThrGlnSer 944
Qy 2794 CTGGCCAAAGATATTTGGGATTTGGAACCTGATTCATGATACAGACCCCAAGCTCA 2853
Db 945 LeuSerThrThrLeuGlyIleAspTyrGluLeu-----AlaProGlnSerSer 960
Qy 2854 GGACAGGTAGAGAGTGAATAGAACCACTTAAGAGACCCCTTACTAAATTCAGCCGGAG 2913
Db 961 GlyLysVal---LysMetAsnHisThrLeuGlnThrThrLeuAlaLysLeuPheGlnGlu 979
Qy 2914 ACTGGCGTTAATGATGATGCTCTCTGCGCCCTTGTGCTTTTATAGGTTAGGAACACC 2973
Db 980 MetHis---GluSerTrpValLysMetLeu---LeuSerLeuLeuArgValArgAla 997
Qy 2974 CCTGGA---CAGTTTGGCTGACCCCTATGAAATTTACTCTACGGGGGACCCGCCCATG 3030
Db 998 ProSerCysSerLeuArgLeuSerSerGlnMetIleTyrGlnArgPro----- 1014
Qy 3031 GTAGAAATGCTTCTGTACATAGTGTGCTGCTGCTTTCCAGCCCTTTGTTCTPAGG 3090
Db 1015 -----PheLeuThrThrAspLeuLeuGlyGluGluLeuHisLysLys 1029
Qy 3091 CTCAGGCACTTGAGTGGTGAGACAAGAGCGGTGGAGGCACTCCGGGAGCGCTACTCA 3150
Db 1030 Leu-----GlnTyrArgPheArgThrAsp 1037
Qy 3151 GGAGGAGGAGACTTGCAGATC-----CCACATCGT----- 3180
Db 1038 SerLysGlyAspPheArgIleTrpLysGlnAsnProAlaSerProHis**GlyTrpGlu 1057
Qy 3181 -----TTCCAAGTG-----GGAGATTTCAGTCTACGTTAGACGCCACCGTGAGGAAC 3228
Db 1058 AsnProPheGlnValAsnProGlyAspGlnValLeuLeuLysAla**ArgAlaGlySer 1077
Qy 3229 CTCGAG-----ACTCGGTGAAGGCGCTTATCTCGTACTTTTGGACACACCA 3276
Db 1078 ProGluAspHisProLeuLeuLysTrpGluGlyProCysTrpValIleLeuThrThrPro 1097
Qy 3277 ACGGCTGTGAAGTCGAGGAATCTCCACCTGGATCCATGCATCCACGTTAAACCGCG 3336
Db 1098 ThrAlaAlaAsnSerGlnGlyIleThrSerTrpValHisLeuSerArgSerGluMetLeu 1117
Qy 3337 CCACCT-----CCGATTTCGGGGTGGGAAGCCGAAAGACGCTGAAAT 3378
Db 1118 SerProLysCysPheGlnThrArgProAspGlyProPheTyrSer**LysProValGlu 1137
Qy 3379 CCCTTTAAGCTTCG 3393
|||||
|||||
```

```
Db 1138 AspLeuLysPheArg 1142
RESULT 3
US-11-042-988-12
; Sequence 12, Application US/11042988
; Publication No. US20050244818A1
; GENERAL INFORMATION:
; APPLICANT: SILICIANO, ROBERT
; APPLICANT: ZHANG, HAILI
; APPLICANT: ZHOU, YAN
; TITLE OF INVENTION: SINGLE CELL ANALYSIS OF HIV REPLICATION CAPACITY AND
; TITLE OF INVENTION: DRUG RESISTANCE
; FILE REFERENCE: 62760(71699)
; CURRENT FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: US/11/042,988
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 12
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-042-988-12
Alignment Scores:
Pred. No.: 3,9e-26 Length: 912
Score: 431.00 Matches: 209
Percent Similarity: 35.9% Conservative: 145
Best Local Similarity: 21.2% Mismatches: 345
Query Match: 7.0% Indels: 286
DB: 39 Gaps: 39
US-10-723-552-3_COPY_2307_5741 (1-3435) x US-11-042-988-12 (1-912)
Qy 85 CTGGTCATACCTGAGTGGCCAGCACCCCTCTTAGGTAGAGACTTATTGACCAAGATCGGA 144
Db 40 ValLeuValGlyProThrProValAsnIleGlyArgAsnLeuLeuThrGlnIleGly 59
Qy 145 GCACAAATTTCTTTTGACAAAGGGGAAACAGAGTGTCTGCAATAACAACCTATCACT 204
Db 60 CysThrLeuAsnPhe-----ProIle--- 66
Qy 205 GTGTTGACCCCTCCAATTAGATGACGAATATCGATATCTCTCCCTTAGTAAAGCCTGAT 264
Db 67 -----SerProIle----- 69
Qy 265 CAAATATACAAATTCTGTTGGAAACAGTTTCCCAAGCCTGGGCGAGAAACCGCAGGGATG 324
Db 69 ----- 69
Qy 325 GGTTCGCAAGCAAGTTCCTCCCAAGTTATTCACTGAAGGCCAGTGCCACACCATG 384
Db 70 -----GluThrValPro-----ValLysLeuLysProGlyMetAspGlyPro 83
Qy 385 TCAGTCAGACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCGCATGTCAA 444
Db 84 LysValLysGlnTrpProLeuThrGluGluLysIleLysAlaLeuValGluIleCysThr 103
Qy 445 AGATTAAATCCACAGGGGCATCTTAGTTCCTGTC-----CAATCTCCTCGAATACCTCCC 498
Db 104 GluMetGluLysGluGlyLysIleSerLysIleGlyProGluAsnProTyrAsnThrPro 123
Qy 499 CTGCTACCGGTTAGAAGCCTGGGACTAATGACTATCGACCATACAGGACTTGAGAGAG 558
Db 124 ValPheAlaIleLysLysLysAspSerThrLysTrpArgLysLeuValAspPheArgGlu 143
Qy 559 GTCAATAAAGCGGTGAGGAT-----ATACACCAACAGCTCCGCAACCTTATAAC 609
Db 144 LeuAsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAlaGly 163
Qy 610 CTCTTGTTGCTCTCTCCACCCCAACGAGCTGGTATATACAGTATTGGAATTAAGGATGCC 669
|||||
|||||
```

[illegible]

1663	QY	---	GAGACTGATGAAC	CAGTGTGACTCAT	GTGATGGCATCA	CTATTGATGTGAGGAGACTGGG	1719
480	Db	PheVal	AsnThrPro	ProLeuVal	LeuLeuTyrGln	LeuGluLysGlu---	496
1720	QY	GTCCGCAAG	ACCTTACAGAC	ATACCGCTG	GAGAGTGTAA	CCTCGTTCACTGCATGAC	1779
497	Db	---	---	ProIle	ValGly---	AlaGluThrPheTyrValAsp	507
1780	QY	GGAAAGCAGC	---	---	TATGTGTGG	AAAGTAAAGAC	1809
508	Db	GlyAla	AlaAsnArg	GluThrLys	LeuGlyLys	AlaGlyTyrValThrAsnArgGlyArg	527
1810	QY	ATGCGTGGG	GGGGTGTGG	AGCGGACCG	CACAGTCTGGG	CCAGCACCTGCCGGA	1869
528	Db	GlnLys	ValValThr	LeuThrAsp	---	---	535
1870	QY	GGAACTTC	CAGCACAAA	AGGCTGAGCTCAT	GGCCCTCACG	CAAGCTTTGCGGTGGCGGAA	1929
536	Db	---	ThrThrAsn	GlnLysThr	GluLeuGln	AlaIleTyrLeuAlaLeuGln---	553
1930	QY	GGGAAATCC	ATAACATTTAT	ACGGACAG	CAGGTATGCTTT	TGCGAGCTGCACACGTACAT	1989
554	Db	GlyLeu	GluValAla	AsnIleValThr	AspSerGln	TyrAlaLeuGly---	568
1990	QY	GGGGCCAT	CTATAAACA	AGGGGTTCCTTAC	CTCAGCAGG	AGGAAATAAAGAACAA	2049
569	Db	---	---	IleIle	GlnAlaGln	ProAspGlnSerGluSer	579
2050	QY	GAGAAATTC	TAAAGCTATT	AGAAGCCGTACAT	TATACCA	AAAAAGGCTAGCTATTATACAC	2109
580	Db	GluLeu	ValAsnGln	IleIleGlu	GlnLeuIle	LysLysValTyrLeuAlaTyr	599
2110	QY	TGTCCTG	CAGATCAGAA	AGCTAAGATCTCAT	TATCCAGAG	AAACACAGATGGCTGACCGG	2169
600	Db	ValPro	AlaHisLys	GlyIle-	-----	GlyGlnGluGlnValAspLys	614
2170	QY	GTTCGCAAG	CAGCGCCG	CGGGTGTAACTCT	TGCTATAAT	TAGAAATGCCCAAGCC	2229
615	Db	LeuVal	SerAlaGly	IleArg	LysValLeu	PheLeu-	626
2230	QY	CCAGAAC	CCAGACAGAC	AGTACAC	CCCTAGAGACTGG	CAAGATATAAAGATAGACCAG	2289
626	Db	---	---	---	---	---	626
2290	QY	TTCTCTG	AGACTCCGGA	AGGAGCCTGCTAT	ACCTCAGATGG	AAGAAATCTGCCCCAC	2349
626	Db	---	---	---	---	---	626
2350	QY	AAAGAAG	GTTT	PAGATATGTCC	ACAGATACATG	TCTAACCCACTAGGAATAAACAC	2409
627	Db	---	AspGly	IleAspLys	AlaGlnAsp	GluHisGluLysTyrHisSerAsnTyrArgAla	645
2410	QY	CTGACAG	AGTGGTCAG	AACATCCC	CTTATCATGTTCT	TGAGGCTACACAGAGTGGCTGAC	2469
646	Db	Met	-----	AlaSer	AspPheAsnLeu	-----ProProValValAlaLys	658
2470	QY	TCGGTGT	CAAAACAT	TGTGTGCC	CTGCGAGCTGGT	TAACTGCTAATCTCTCCAGAAATGCCT	2529
659	Db	GluIle	ValAlaSer	CysAsp	LysCysGlnLeu	---	669
2530	QY	CCAGGAAG	AGAGACTA	AGGGGAGC	-----	CACCGAGCGCTACTCGGAAGTGGAC	2580
670	Db	LysGly	GluAlaMet	HisGlyGln	ValAspCys	ProGlyIle---TyrGlnLeuAsp	688
2581	QY	TTCTACT	GAGGTAA	ACCGGCTAA	TATACGGA	AAACAAATACCTATTGGTTTTGTAGACACC	2640
689	Db	CysThr	HisLeu	Glu-	-----	GlyLysValIleLeuValAlaValHisVal	703
2641	QY	TTTTCA	GATGGGTAG	AGGCTTATCT	TACTAAG	AAAGAGACTTCAACCGTGGTGGCTAAA	2700
704	Db	AlaSer	GlyTyrIle	GluAlaGlu	ValIlePro	AlaGluThrGlyGlnGluThrAlaTyr	723


```
QY 1717 GGGGTCGCGAAGGACCTTTACAGACATACCGCTGACTGGAGAAGTGCTAACTGGTTCACT 1776
Db 435 -----ProleValGly-----AlaGluThrPheTyrVal 444
QY 1777 GACGGAAGCAGC-----TATGTGGTGGAAAGGTAAG 1806
Db 445 AspGlyAlaAAsnArgGluThrLysLeuGlyLysAlaGlyTyrValThrAsnArgGly 464
QY 1807 AGGATGGCTGGGCGCGGTGGTGGACCGGACCGCCACGATCTGGGCCAGCAGCTGCCG 1866
Db 465 ArgGlnLysValValThrLeuThrAsp-----473
QY 1867 GAAGGAAGCTTCAGACAAAGCGCTCATGGCCCTACGCAAGCTTTCGGCTGGGCC 1926
Db 474 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyrLeuAlaLeuGln---Asp 490
QY 1927 GAAGGGAATCCATAAACATTATACGACGACGAGGTATGCTTTGGCACTGCACAGTA 1986
Db 491 SerGlyLeuGluValAsnIleValThrAspSerGlnTyrAlaLeuGly-----506
QY 1987 CATGGGCGCATCTATAAACAAAGGGGGTGTCTTACCTCAGCAGGAGGGAATAAAGAAC 2046
Db 507 -----IleLeuGlnAlaGlnProAspGlnSerGlu 516
QY 2047 AAAGAGGAATTTAAGCCTATTAGAAGCCGTACATTTTACCAGAAAGGCTAGCTATTATA 2106
Db 517 SerGluLeuValAsnGlnIleLeuGlnLeuLysLysGluLysValTyrLeuAla 536
QY 2107 CACTGCTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAACAGATGGCTGAC 2166
Db 537 TrpValProAlaHisLysGlyLe-----GlyGlyAsnGluGlnValAsp 551
QY 2167 CGGGTGGCAAGCAGGACCCAGGGTGT 2196
Db 552 LysLeuValSerAlaGlyIleArgLysVal 561

RESULT 5
US-10-507-928-12
; Sequence 12, Application US/10507928
; Publication No. US2005026024A1
; GENERAL INFORMATION:
; APPLICANT: POWDERMED LIMITED AND GLAXO GROUP LIMITED
; TITLE OF INVENTION: ADJUVANT
; FILE REFERENCE: N.88232B GCW
; CURRENT APPLICATION NUMBER: US/10/507,928
; CURRENT FILING DATE: 2004-09-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of the coding insert in p73i-RT3
US-10-507-928-12

Alignment Scores:
Pred. No.: 5,07e-16 Length: 562
Score: 301.00 Matches: 145
Percent Similarity: 37.2% Conservative: 107
Best Local Similarity: 21.6% Mismatches: 264
Query Match: 4.9% Indels: 154
Db: 6 Gaps: 26

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-10-507-928-12 (1-562)
QY 322 ATGGGTTTGGCAAGCAAGTCCCCACACAGTTATTTCACCTGAAGCCAGCGGTGCACACCA 381
Db 1 MetGlyProIleSerProIleThrValSerValLysLeuLysLeuLysProGlyMetAspGly 20
QY 382 GTGTGAGTACAGACGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCGCATGTC 441
Db 21 ProlLysValLysGlnTrpProLeuThrGluGluLysLysLysAlaLeuValGluIleCys 40
```

```
QY 442 CAAGATTAAATCCAAACAGGCGATCCTAGTTCTCTGTC-----CAATCTCCCTGGAATACT 495
Db 41 ThrGluMetGluLysGluGlyLysIleSerLysLysIleGlyProGluAsnProTyrAsnThr 60
QY 496 CCCCTGCTACCGGTTAGAAAGCCCTGGGACTAATGACTATCGACGAGTACAGCACTTGAGA 555
Db 61 ProValPheAlaIleLysLysLysAspSerThrLysIleArgLysLeuValAspPheArg 80
QY 556 GAGTCTAATAAACGGGTGCGAGAT-----ATACACCCCAACAGCTCCCAACCCCTTAT 606
Db 81 GluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAla 100
QY 607 AACCTCTTGTGTCTCTCCACCCCAACGGAGCTGTATACAGTATTGGATTGATAAAGGAT 666
Db 101 GlyLeu-----LysLysLysLysSerValThrValLeuAspValGlyAsp 115
QY 667 GCCTTCTTCTGCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTCGAATGGAGA 726
Db 116 AlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIlePro 135
QY 727 GATCCAGGTACGGGAAGAACCGGG---CAGCTCACCTGGACCCGACCTGCCCAAGGGTTC 783
Db 136 SerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGlyTrp 155
QY 784 AAGAACTCCCGCAGCATCTTTTGACGAAGCCCTACACAGAGACCTGGCCCACTTCAGGATC 843
Db 156 GlySerProAlaIlePheGlnSerSerMetThrLysIleLeuGluProPheArgLys 175
QY 844 CAACACCTCCTCAGGTGACCTCTCCTCAGTACGTGATGACCTCTTCTGGCGGAGCCACC 903
Db 176 GlnAsnProAspIleValIleTyrGlnTyrMetAspAspLeuTyrVal---GlySerAsp 194
QY 904 AAACAGGACTCTTAGAAGGACCAAG-----GCATCTACTGCTGGAAATTTGCTGACCTA 957
Db 195 LeuGluIleGlyGlnHisArgThrLysIleGluLeuArgGlnHisLeuLeuArgTrp 214
QY 958 GGCTACAGAGCTCTGTAGAAAGCCAGATTTCAGGAGAGAGGTAACTACTTG---1014
Db 215 GlyLeuThrThrProAspLysLysHisGln-----LysGluProProPheLeuTrp 231
QY 1015 ---GGGTACAGTTTTCGGGACGGCGACGCTGCTGACGGAGGACGGAAGAAACTGTA 1071
Db 232 MetGlyTyrGluLeuHisPro---AspLysTrp-----ThrVal 243
QY 1072 GTCCAGATACCGGCCCCCAACC-----ACGCCAAACAAATAGAGAGTTTGTG 1119
Db 244 GlnProIleValLeuProGluLysAspSerTrpThrValAsnAspIleGlnLysLeuVal 263
QY 1120 GGGACAGCTGGATTTCAGACTGTGGATCCCGGGGTTT-----1158
Db 264 GlyLysLeuAsnTrpAlaSerGlnIleTyrProGlyIleLysValArgGlnLeuCysLys 283
QY 1159 -----GCGACCTTAGCAGCCCCCACTCTACCCGCTAACCAAA-----1194
Db 284 LeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGluGluAlaGluLeu 303
QY 1195 -----GAAAAGGGGAAATTCCTGGGCTCTCTGAGCACCAGAGGCAATTCATGCT 1245
Db 304 GluLeuAlaGluAsnArgGluIleLysGluProValHisGlyValTyrTyrAspPro 323
QY 1246 ATCAAAAAGGCCCTCTGTAGCGCACCTGTCTGGCCCTCCCTGACGTAACCTAAACCCCTT 1305
Db 324 SerLysAspLeuIleAlaGluIleGlnLysGlnGlyGlnTrpThr-----Tyr 341
QY 1306 ACCCTTTATGTGATAGCGTAAGGAGTAGCCCGGGAGGATTTTAAACCCCAACCTAGGA 1365
Db 342 GlnIleTyrGlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArgMetArgGly 361
QY 1366 CATGGAGAGACCTGTGCTGATCAAGAGCTCGATCCCTGTAGCAGCAGGTGGTGG 1425
Db 362 AlaHisThrAsnAspValLysGlnLeuThrGluAlaValGlnLysIleThrThr-----379
```

```
QY 1426 CCCATATGCTGAAGGCTATCGCAGCTGTGGCCTACTGTCAGGAGCGCTGACAAATTG 1485
Db 380 -----GluserialletrpGlyLys----- 387
QY 1486 ACTTTGGACAGATATACTGTAATAGCCCCCATGTCATGGAGAACATGTCGGCAG 1545
Db 388 -----ThrProlysPheLysLeuProilleGlnLysGlu 398
QY 1546 CCCCAGACCGATGATGACCAACGCCGCGCATGACCCACTATCAAGCCTGCTTCTACA 1605
Db 399 ThrTrpGluThrTrpTrpThrGlu-----TyrTrpGln----- 409
QY 1606 GAGAGGTCAGCTTCGCTCCACAGCCGCTCTCAACCCCTGCCACTCTCTCTGCTGAA--- 1662
Db 410 -----AlaThrTrpIleProGluTrp 416
QY 1663 -----GAGACTGTAACAGTCACTCATGATTCGCTCACTCAATATTGATGAGGAGACT 1716
Db 417 GluPheValAsnThrProProLeuValLysLeuTrpThrGlnLeuGluLysGlu----- 434
QY 1717 GGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAGTCTTAACCTGGTTCAC 1776
Db 435 -----ProilleValGly---AlaGluThrPheTyrVal 444
QY 1777 CACGGAACGACG-----TATGTGTGGAAGGTAAAG 1806
Db 445 AspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGlyTyrValThrAsnArgGly 464
QY 1807 AGGATGCTGGCGGGGCTGGAGCGGAGCCCGCAGCATCTGGGCCAGCAGCTGCCG 1866
Db 465 ArgGlnLysValValThrLeuThrAsp----- 473
QY 1867 GAAGGAACCTCAGCACAAAGGCTGAGCTCATGCGCTTCAGCAAGCTTTTCGGCTGCC 1926
Db 474 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyrLeuAlaLeuGln---Asp 490
QY 1927 GAAGGAAATCCATAACATTTATACGACAGAGGTATGCTTTGGCAGCTGCACAGTA 1986
Db 491 SerGlyLeuGluValAsnIleValThrAspSerGlnTyrAlaLeuGly----- 506
QY 1987 CATGGGCGCATCTATAACAAAGGGGTGTCTTACCTCAGCAGGAGGGAATAAGAAC 2046
Db 507 -----IleleGlnAlaGlnProAspGlnSerGlu 516
QY 2047 AAAGAGGAATCTTAAGCTTATTAGAAGCGCTACATTTTACCAGAAAGGCTAGCTATTATA 2106
Db 517 SerGluLeuValAsnGlnIleleGluLeuLysLysGlyLysValTyrLeuAla 536
QY 2107 CACTGCTCTGACATCAGAAAGCTTAAGATCTCATATCCAGAGAAACAGATGGCTGAC 2166
Db 537 TrpValProAlaHisLysGlyIle-----GlyGlyAsnGluGlnValAsp 551
QY 2167 CGGCTGCCAGCAGCGGAGCCCGGGTGT 2196
Db 552 LysLeuValSerAlaGlyIleArgLysVal 561
```

RESULT 6

```
US-11-029-465-10
; Sequence 10, Application US/11029465
; Publication No. US20050256070A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Ralph P.
; APPLICANT: Thomsen, Lindy
; APPLICANT: Van-Wely, Catherine
; APPLICANT: Ertl, Peter
; TITLE OF INVENTION: Adjuvant
; FILE REFERENCE: 033267-015
; CURRENT FILING DATE: 2005-01-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 562
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of RT insert of p7077-RT3
US-11-029-465-10

Alignment Scores:
Pred. No.: 5,07e-16 Length: 562
Score: 301.00 Matches: 145
Percent Similarity: 37.2% Conservative: 104
Best Local Similarity: 21.6% Mismatches: 267
Query Match: 4.9% Indels: 154
DB: 7 Gaps: 26
```

```
US-10-723-552-3_COPY_2307_5741 (1-3435) x US-11-029-465-10 (1-562)
QY 322 ATGGGTTTGGCAAGCAAGTTCCTCCACAGTTATTCACTGAGGCGCAGTGCCACACCA 381
Db 1 MetGlyProIleSerProIleGluThrValSerValLysLeuLysProGlyMetAspGly 20
QY 382 GTGTCAGTCACAGACATACCCCTTCAGTAAAGAGCTCAAGAGGAATTCGGCGCATGTC 441
Db 21 ProLysValLysGlnTrpProLeuThrGluLysLysLysAlaLeuValGluLysCys 40
QY 442 CAAAGATTAATCCAAACAGGCGCATCTAGTTCCTGTC-----CAATCTCCCTGGAATACT 495
Db 41 ThrGluMetGluLysGluGlyLysLysLysLysLysLysLysLysLysLysLysLysLys 60
QY 496 CCCCTGCTACCGGTAGAAAGCCCTGGAGCACTAATCACTATCGACAGTACAGACTTGAGA 555
Db 61 ProValPheAlaIleLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 80
QY 556 GAGGTCAATAAAGCGGTGCAGGAT-----ATACACCCCAACAGTCCCAACCCCTTAT 606
Db 81 GluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAla 100
QY 607 AACCTCTGTGTGCTCTCCCAACCGAGAGTGGTATACAGTATTGGACTTAAAGAT 666
Db 101 GlyLeu-----LysLysLysLysSerValThrValLeuAspValGlyAsp 115
QY 667 GCCTTCTTCCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTGTGCTTCGAATGGAGA 726
Db 116 AlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIlePro 135
QY 727 GATCCAGGTACGGGAAGAACCGGG---CAGCTCACCTGCGCCGCTGCTCCCAAGGGTTC 783
Db 136 SerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGlyTrp 155
QY 784 AAGAACTCCCGACCATCTTTTGACGAAGCCCTACACAGAGACCTTGGCCCACTTCAGGATC 843
Db 156 LysGlySerProAlaIlePheGlnSerSerMetThrLysLysLysLysLysLysLysLys 175
QY 844 CAACACCTCAGGTGACCTCTCCAGTACGTGGATGACCTGCTTCTGGCGGGAGCCACC 903
Db 176 GlnAsnProAspIleValIleTyrGlnTyrMetAspAspLeuTyrVal---GlySerAsp 194
QY 904 AAACAGGACTGCTTAGAAGGCGCAAG---GCCTACTGCTGGAATGTCTGACCTA 957
Db 195 LeuGluIleGlyGlnHisArgThrLysIleGluLeuLeuArgGlnHisLeuLeuArgTrp 214
QY 958 GGCTACAGACCTCTGCTAAGAAAGCGCCAGATTTGCAGGAGAGAGAGTAAACATCTTG--- 1014
Db 215 GlyLeuThrThrProAspLysLysHisGln-----LysGluProProPheLeuTrp 231
QY 1015 ---GGGTACAGTTTGGGAGCGGCGACCGATGCTGACGGAGGACCGAAGAAACATCTGTA 1071
Db 232 MetGlyTyrGluLeuHisPro---AspLysTrp-----ThrVal 243
QY 1072 GTCAGATACCGGCCCAACC-----ACAGCCAAACAAATGAGAGAGTCTTTTG 1119
Db 244 GlnProIleValLeuProGluLysAspSerTrpThrValAsnAspIleGlnLysLeuVal 263
QY 1120 GGGACAGCTGGATTTTTCAGACTGTGATCCCGGGTGT----- 1158
```

```
Db      264 GlyLysLeuAsnTrpAlaSerGlnIleTyrProGlnIleLysValArgGlnLeuCysLys 283
      |||      :::      |||||
QY      1159 -----GCGACCTTAGCAGCCCACTCTACCCGCTAACCAAA----- 1194
      |||      :::      |||||
Db      284 LeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGluGluAlaGluLeu 303
      |||      |||      |||
QY      1195 -----GAAAGGGGAATTCCTGGGCTCTGAGCACCAGAGCGCATTTGATGCT 1245
      |||      |||      |||
Db      304 GluLeuAlaGluAsnArgGluLeuLysGluProValHisGlyValTyrTyrAspPro 323
      |||      |||      |||
QY      1246 ATCAAAAGGCCCTCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACTAAACCCCTT 1305
      |||      |||      |||
Db      324 SerLysAspLeuIleAlaGluIleGlnLysGlnGlyGlnIleTyrThr-----Tyr 341
      |||      |||      |||
QY      1306 ACCCTTTATGTGATGAGCGTAAGGAGTAGCCCGGGAGTTTTAACCCAAACCCCTAGGA 1365
      |||      |||      |||
Db      342 GlnIleTyrGlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArgMetArgGly 361
      |||      |||      |||
QY      1366 CCATGGAGAACCTGTCCCTACCTGTCAAAGAGCTCGATCTGTAGCAGCTGGTGG 1425
      |||      |||      |||
Db      362 AlaHisThrAsnAspValLysGlnLeuThrGluAlaValGlnLysIleThr----- 379
      |||      |||      |||
QY      1426 CCATATGCTCGTAGGCTATGCGAGCTGTGGCCATCTGCTCAAGGACGCTGACAAATTG 1485
      |||      |||      |||
Db      380 -----GluSerIleValIleTyrGlyLys----- 387
      |||      |||      |||
QY      1486 ACTTTGGGACAGATATACTGTATAAGCCCCCATGCACTATGAGAAACATCGTTCGGCAG 1545
      |||      |||      |||
Db      388 -----ThrProLysPheLysLeuProIleGlnLysGlu 398
      |||      |||      |||
QY      1546 CCCCAGACCGATGATGATACCAACGCCCGCATGCCCACTATCAAAAGCTGCTTCACA 1605
      |||      |||      |||
Db      399 ThrTrpGluThrTrpTrpThrGlu-----TyrTrpGln----- 409
      |||      |||      |||
QY      1606 GAGAGGTCACGTTTCGCTCCACCGCGCTCTCAACCCCTGCCACTCTTCGCTCAA--- 1662
      |||      |||      |||
Db      410 -----AlaThrTrpIleProGluTrp 416
      |||      |||      |||
QY      1663 -----GAGACTGATGAACCACTGATCATGATTCGCATCACTATGATGAGGAGACT 1716
      |||      |||      |||
Db      417 GluPheValAsnThrProProLeuValLysLeuTyrTrpGlnLeuGluLysGlu----- 434
      |||      |||      |||
QY      1717 GGGGTCGCGAAGACCTTACAGACATACGCTGACTGGAGAAGTCTTAACCTGGTTCACT 1776
      |||      |||      |||
Db      435 -----ProfileValGly-----AlaGluThrPheTyrVal 444
      |||      |||      |||
QY      1777 GACGGAAGCAGC-----TATGTGTGGAGGTAAG 1806
      |||      |||      |||
Db      445 AspGlyAlaIleAsnArgGluThrLysLeuGlyLysAlaGlyTyrValThrAsnArgGly 464
      |||      |||      |||
QY      1807 AGGATGGCTGGGCGCGTGGTGAGCGGACCCGACGATCTGGGCGCAGCAGCCTGCGC 1866
      |||      |||      |||
Db      465 ArgGlnLysValValThrLeuThrAsp----- 473
      |||      |||      |||
QY      1867 GNAGAACTTCAGCAAAAGCTCAGCTCATGCGCCCTCACCAAGCTTTGGCGTGCC 1926
      |||      |||      |||
Db      474 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyrLeuAlaLeuGln---Asp 490
      |||      |||      |||
QY      1927 GAAGGGAATCCATAAACTTATACGGACAGCATGTCCTTTGGCAGTGCACACGTA 1986
      |||      |||      |||
Db      491 SerGlyLeuGluValAsnIleValThrAspSerGlnTyrAlaLeuGly----- 506
      |||      |||      |||
QY      1987 CATGGGCGCATCTATAAAAGGGGGTTGCTTACCTCAGCAGGGGGAATAAAGAAC 2046
      |||      |||      |||
Db      507 -----IleIleGlnAlaGlnProAspGlnSerGlu 516
      |||      |||      |||
QY      2047 AAAGAGGAATTCATAGCCTATTAGACCGCTACATTTACCAAAAGGCTAGCTATTATA 2106
      |||      |||      |||
Db      517 SerGluLeuValAsnGlnIleIleGluGlnLeuIleLysLysGluLysValTyrLeuAla 536
      |||      |||      |||
QY      2107 CACTGTCTGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAACCCAGATGCTGAC 2166
      |||      |||      |||
```

```
Db      537 TrpValProAlaHisLysGlyIle-----GlyGlyAsnGluGlnValAsp 551
      |||      |||      |||
QY      2167 CGGTTTGGCAAGCAGGAGCCCGGCTGTT 2196
      |||      |||      |||
Db      552 LysLeuValSerAlaGlyIleArgLysVal 561
      |||      |||      |||
RESULT 7
US-11-029-465-12
; Sequence 12, Application US/11029465
; Publication No. US20050256070A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Ralph P.
; APPLICANT: Thomsen, Lindy
; APPLICANT: Van-Wely, Catherine
; APPLICANT: Ertl, Peter
; TITLE OF INVENTION: Adjuvant
; FILE REFERENCE: 033267-015
; CURRENT APPLICATION NUMBER: US/11/029,465
; CURRENT FILING DATE: 2005-01-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of the coding insert in
; OTHER INFORMATION: p731-RT3
US-11-029-465-12
Alignment Scores:
Pred. No.:      5,76e-15      Length:      561
Score:          287.50      Matches:     139
Percent Similarity: 36.1%      Conservative: 105
Best Local Similarity: 20.6%      Mismatches: 265
Query Match:      4.6%      Indels:      167
DB:              7          Gaps:          24
US-10-723-552-3_COPY_2307_5741 (1-3435) x US-11-029-465-12 (1-561)
QY      322 ATGGGTTGGCAAGAGTTCCCCCAACAAGTTATTCAACTGGAAGCGCAGTGCCACACA 381
      |||      |||      |||
Db      1 MetGlyProIleSerProIleGluThrValSerValLysLeuLysProGlyMetAspGly 20
      |||      |||      |||
QY      382 GTGTCAGTCAGACAGTACCCCTTGAGTAAACAAGCTCAAGAGGAAATTCGCGCGCATGTC 441
      |||      |||      |||
Db      21 ProLysValLysGln---ProLeuThrGluGluLysIleLysAlaLeuValGluLeuCys 39
      |||      |||      |||
QY      442 CAAAGATTAAATCCCAACAGGCGATCCTAGTTCCTGTC-----CAATCTCCTCGAATACT 495
      |||      |||      |||
Db      40 ThrGluMetGluLysGluGlyLysIleSerLysIleGlyProGluAsnProTyrAsnThr 59
      |||      |||      |||
QY      496 CCCCTGTCTACCGGTTAGAAAGCCTGGGACTAATGATATCGACACCACTGACAGCTTGAGA 555
      |||      |||      |||
Db      60 ProValPheAlaIleLysLysLysAspSerThrLysLysLysLysLysLysLysLysLys 79
      |||      |||      |||
QY      556 GAGGTCAATAAAGCGTGCAGAT-----ATACACCCCAACAGTCCCGAACCCCTTAT 606
      |||      |||      |||
Db      80 GluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAla 99
      |||      |||      |||
QY      607 AACCTCTTGTGCTCTCCACCCCAACGAGCTGTATACAGTATTGCACTTAAGGAT 666
      |||      |||      |||
Db      100 GlyLeu-----LysLysLysLysSerValThrValLeuAspValGlyAsp 114
      |||      |||      |||
QY      667 GCCTTCTTCTGCTGAGATTACACCCCACTAGCCCAACCACTTTTTCCTTGAATGGAGA 726
      |||      |||      |||
Db      115 AlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIlePro 134
      |||      |||      |||
QY      727 GATCAGGTACGGGAGAACCGGG---CAGCTCAGCTGAGCCCGGCTCCCGCAAGGTC 783
      |||      |||      |||
Db      135 SerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGlyTrp 154
      |||      |||      |||
QY      784 AAGAACTCCCGCACCATCTTTGACGAAGCCCTTACACAGAGACCTGGCCAACTTCAGGATC 843
```


Db 155 LysGlySerProAlaIlePheGlnSerMetThrLysIleLeuGluProPheArgLys 174
||| ||||| ||||| : : : : : ||| |||||
Qy 844 CAACACCTCAGGTGACCTCTCCAGTACGTGGATGACCTCTTCGCGCGGACCCACC 903
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 175 GlnAsnProAspIleValIleTyrGlnTyrMetAspLeuTyrValGly 191
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 904 AAACAGGACTGCTAGAGGACGAGGACACTACTGCTGGAATGCTGACCTAGGCTAC 963
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 192 -----SerAspLeuGluIle 196
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 964 AGAGCCTCTGTAAGAGCCCGAGATTTGCAGGAGAGGTAAACATCTTGGGTACAGT 1023
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 197 GlyGlnHisThrArgLysIleGluIleuArgGlnHisLeuLeuArgTyrGlyLeuThr 216
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1024 TTGGGGAGCGGCGACGCA 1053
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 217 ThrProAspLysLysHisGlnLysGluProPheLeuTyrMetGlyTyrGluLeuHis 236
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1054 GCACGGAAGAACTGTAGTCCAGATACCGGCCCAACC-----ACAGCCAAA 1101
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 237 ProAspLysTyrThrValGlnProIleValLeuProGluLysAspSerTyrThrValAsn 256
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1102 CABATGAGAGATTTTGGGACACCTGATTTTGCAGCTGTGGATCCCGGGT--- 1158
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 257 AspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSerGlnIleTyrProGlyIleLys 276
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1159 -----CGCAGCTTAGCAGCCCACTCTACCCGCTA 1188
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 277 ValArgGlnLeuCysLysLeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeu 296
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1189 ACCAAA-----GAAAAGGGGAATTCCTCGGCTCTGAGCAC 1227
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 297 ThrGluGluAlaGluLeuAlaGluAsnArgGluIleLeuLysGluProValHis 316
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1228 CAGAAGGCAATTGATGCTATCAAAAGCCCTGTGAGCGCACCTGTCTGGCCCTCCCT 1287
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 317 GlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGlnGly 336
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1288 GACGTAACCTAACCTTACCTTTATGTGTGATGAGCGTAAAGGAGTAGCCCGGGAGTT 1347
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 337 GlnTrpThr-----TyrGlnIleTyrGlnLysAsnLeuLysThrGlyLys 354
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1348 TTAACCCAAACCTTAGGACCATGGAAGACCTGTGCTGCTCAAAAGAGCTCGAT 1407
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 355 TyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeuThrGluAlaValGln 374
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1408 CCTGTAGCCAGTGTGGCCCATATGCTGAGGCTATCGCAGCTGTGGCCATCTGGTC 1467
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 375 LysIleThr-----GluSerIleValIleTrpGly 385
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1468 AAGGACCTGACAAATTGCTTTGGGACAGAAATATACTGTAATAGCCCCCATGCTG 1527
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 386 Lys-----ThrProLysPheLys 391
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1528 GAGAACATCGTTGGCAGCCCCCAGACCGATGGATGATGATGATGATGATGATGAT 1587
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 392 LeuProIleGlnLysGluThrTrpGluThrTrpTrpGlu-----TyrTrp 407
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1588 CAAAGCTGCTTCTCAGAGAGGGTCACTGCTCCACGAGCGGCTCTCAACCTGCC 1647
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 408 Gln-----Ala 409
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1648 ACTCTTCTGCTGAA-----GAGACTGATGAACGAGTCACTCATGATGATGATGAT 1698
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 410 ThrTrpIleProGluTrpGluPheValAsnThrProLeuValLysLeuTrpTyrGln 429
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1699 CTATGATGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAA 1758
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 430 LeuGluLysGlu-----ProIleValGly--- 437
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1759 GTGCTAACCTGGTTCACTGACGAGGAGC----- 1788
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 438 AlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGly 457
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1789 TATCTGTGGAAAGGTAAAGAGATGGCTGGGCGCGGTGGTGGAGCGGACCGCACGATC 1848
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 458 TyrValThrAsnArgGlyArgGlnLysValValThrLeuThrAsp----- 472
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1849 TGGCCAGCAGCCTGCCGGAAGGAACTTCAGACACAAAGGCTGAGCTCATGCGCTCAGC 1908
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 473 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyr 484
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1909 CAAGCTTTGGCTGGCGGAGGGAATTCATAACATTTATACGACAGCAGGATGCTCC 1968
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 485 LeuAlaLeuGln---AspSerGlyLeuGluValAsnIleValThrAspSerGlnTyrAla 503
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1969 TTTGCGACTGCACACGTACATGGGCGCATCTATAAACAAGGGGGTGTCTTACCTCACA 2028
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 504 LeuGly-----IleIleGlnAla 509
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2029 GGGAGGGAATAAGAAACAAGAGAAATTCCTAAGCCTATTAGAACGGCTGATACCTA 2088
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 510 GlnProAspGlnSerGluSerGluLeuValAsnGlnIleIleGluGlnLeuIleLysLys 529
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2089 AAAAGGCTAGCTATTATACACTGCTCTGGACATCAGAAAGCTAAGATCTCATATCCAGA 2148
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 530 GluLysValTyrLeuAlaTrpValProAlaHisLysGlyIle-----Gly 544
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2149 GGAACACAGATGGCTGACCGGTGGCAAGCAGGCGAGCGGTGT 2196
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 545 GlyAsnGluGlnValAspLysLeuValSerAlaGlyIleArgLysVal 560
||| : : : : : ||||| : : : : : ||||| : : : : :
RESULT 8
US-11-022-562-221
; Sequence 221, Application US/11022562
; Publication No. US20050249742A1
; GENERAL INFORMATION:
; APPLICANT: Ruprecht, Ruth M.
; APPLICANT: Shisong, Jiang
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; FILE REFERENCE: A CVTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
; CURRENT APPLICATION NUMBER: US/11/022,562
; PRIOR FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: PCT/US03/20322
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392718
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 31, 97, 140, 141, 144, 178
; OTHER INFORMATION: Xaa = Any Amino Acid
US-11-022-562-221
Alignment Scores:
Pred. No.: 1,35e-09 Length: 210
Score: 217.50 Matches: 65
Percent Similarity: 47.9% Conservative: 37
Best Local Similarity: 30.5% Mismatches: 82
Query Match: 3.5% Indels: 29
DB: Gaps: 9
US-10-723-552-3_copy_2307_5741 (1-3435) x US-11-022-562-221 (1-210)
Qy 478 CAATCTCCCTCGAATACTCCCTGCTACCGGTTAGAAAGCGCTGGACTAATGACTATCGA 537
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 16 GluAsnProTyrAsnThrProValPheAlaIleLysLysLysAsp***ThrLysTrpArg 35
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 538 CCAGTACAGGACTTGAGAGAGGTCAATAAAGCGGTGCAGGAT-----ATACACCCA 588
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 155 LysGlySerProAlaIlePheGlnSerMetThrLysIleLeuGluProPheArgLys 174
||| ||||| ||||| : : : : : ||| |||||
Qy 844 CAACACCTCAGGTGACCTCTCCAGTACGTGGATGACCTCTTCGCGCGGACCCACC 903
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 175 GlnAsnProAspIleValIleTyrGlnTyrMetAspLeuTyrValGly 191
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 904 AAACAGGACTGCTAGAGGACGAGGACACTACTGCTGGAATGCTGACCTAGGCTAC 963
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 192 -----SerAspLeuGluIle 196
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 964 AGAGCCTCTGTAAGAGCCCGAGATTTGCAGGAGAGGTAAACATCTTGGGTACAGT 1023
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 197 GlyGlnHisThrArgLysIleGluIleuArgGlnHisLeuLeuArgTyrGlyLeuThr 216
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1024 TTGGGGAGCGGCGACGCA 1053
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 217 ThrProAspLysLysHisGlnLysGluProPheLeuTyrMetGlyTyrGluLeuHis 236
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1054 GCACGGAAGAACTGTAGTCCAGATACCGGCCCAACC-----ACAGCCAAA 1101
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 237 ProAspLysTyrThrValGlnProIleValLeuProGluLysAspSerTyrThrValAsn 256
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1102 CABATGAGAGATTTTGGGACACCTGATTTTGCAGCTGTGGATCCCGGGT--- 1158
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 257 AspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSerGlnIleTyrProGlyIleLys 276
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1159 -----CGCAGCTTAGCAGCCCACTCTACCCGCTA 1188
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 277 ValArgGlnLeuCysLysLeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeu 296
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1189 ACCAAA-----GAAAAGGGGAATTCCTCGGCTCTGAGCAC 1227
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 297 ThrGluGluAlaGluLeuAlaGluAsnArgGluIleLeuLysGluProValHis 316
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1228 CAGAAGGCAATTGATGCTATCAAAAGCCCTGTGAGCGCACCTGTCTGGCCCTCCCT 1287
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 317 GlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGlnGly 336
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1288 GACGTAACCTAACCTTACCTTTATGTGTGATGAGCGTAAAGGAGTAGCCCGGGAGTT 1347
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 337 GlnTrpThr-----TyrGlnIleTyrGlnLysAsnLeuLysThrGlyLys 354
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1348 TTAACCCAAACCTTAGGACCATGGAAGACCTGTGCTGCTCAAAAGAGCTCGAT 1407
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 355 TyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeuThrGluAlaValGln 374
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1408 CCTGTAGCCAGTGTGGCCCATATGCTGAGGCTATCGCAGCTGTGGCCATCTGGTC 1467
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 375 LysIleThr-----GluSerIleValIleTrpGly 385
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1468 AAGGACCTGACAAATTGCTTTGGGACAGAAATATACTGTAATAGCCCCCATGCTG 1527
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 386 Lys-----ThrProLysPheLys 391
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1528 GAGAACATCGTTGGCAGCCCCCAGACCGATGGATGATGATGATGATGATGATGAT 1587
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 392 LeuProIleGlnLysGluThrTrpGluThrTrpTrpGlu-----TyrTrp 407
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1588 CAAAGCTGCTTCTCAGAGAGGGTCACTGCTCCACGAGCGGCTCTCAACCTGCC 1647
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 408 Gln-----Ala 409
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1648 ACTCTTCTGCTGAA-----GAGACTGATGAACGAGTCACTCATGATGATGATGAT 1698
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 410 ThrTrpIleProGluTrpGluPheValAsnThrProLeuValLysLeuTrpTyrGln 429
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1699 CTATGATGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAA 1758
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 430 LeuGluLysGlu-----ProIleValGly--- 437
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1759 GTGCTAACCTGGTTCACTGACGAGGAGC----- 1788
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 438 AlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGly 457
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1789 TATCTGTGGAAAGGTAAAGAGATGGCTGGGCGCGGTGGTGGAGCGGACCGCACGATC 1848
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 458 TyrValThrAsnArgGlyArgGlnLysValValThrLeuThrAsp----- 472
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1849 TGGCCAGCAGCCTGCCGGAAGGAACTTCAGACACAAAGGCTGAGCTCATGCGCTCAGC 1908
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 473 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyr 484
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1909 CAAGCTTTGGCTGGCGGAGGGAATTCATAACATTTATACGACAGCAGGATGCTCC 1968
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 485 LeuAlaLeuGln---AspSerGlyLeuGluValAsnIleValThrAspSerGlnTyrAla 503
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1969 TTTGCGACTGCACACGTACATGGGCGCATCTATAAACAAGGGGGTGTCTTACCTCACA 2028
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 504 LeuGly-----IleIleGlnAla 509
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2029 GGGAGGGAATAAGAAACAAGAGAAATTCCTAAGCCTATTAGAACGGCTGATACCTA 2088
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 510 GlnProAspGlnSerGluSerGluLeuValAsnGlnIleIleGluGlnLeuIleLysLys 529
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2089 AAAAGGCTAGCTATTATACACTGCTCTGGACATCAGAAAGCTAAGATCTCATATCCAGA 2148
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 530 GluLysValTyrLeuAlaTrpValProAlaHisLysGlyIle-----Gly 544
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2149 GGAACACAGATGGCTGACCGGTGGCAAGCAGGCGAGCGGTGT 2196
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 545 GlyAsnGluGlnValAspLysLeuValSerAlaGlyIleArgLysVal 560
||| : : : : : ||||| : : : : : ||||| : : : : :
RESULT 8
US-11-022-562-221
; Sequence 221, Application US/11022562
; Publication No. US20050249742A1
; GENERAL INFORMATION:
; APPLICANT: Ruprecht, Ruth M.
; APPLICANT: Shisong, Jiang
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; FILE REFERENCE: A CVTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
; CURRENT APPLICATION NUMBER: US/11/022,562
; PRIOR FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: PCT/US03/20322
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392718
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 31, 97, 140, 141, 144, 178
; OTHER INFORMATION: Xaa = Any Amino Acid
US-11-022-562-221
Alignment Scores:
Pred. No.: 1,35e-09 Length: 210
Score: 217.50 Matches: 65
Percent Similarity: 47.9% Conservative: 37
Best Local Similarity: 30.5% Mismatches: 82
Query Match: 3.5% Indels: 29
DB: Gaps: 9
US-10-723-552-3_copy_2307_5741 (1-3435) x US-11-022-562-221 (1-210)
Qy 478 CAATCTCCCTCGAATACTCCCTGCTACCGGTTAGAAAGCGCTGGACTAATGACTATCGA 537
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 16 GluAsnProTyrAsnThrProValPheAlaIleLysLysLysAsp***ThrLysTrpArg 35
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 538 CCAGTACAGGACTTGAGAGAGGTCAATAAAGCGGTGCAGGAT-----ATACACCCA 588
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 155 LysGlySerProAlaIlePheGlnSerMetThrLysIleLeuGluProPheArgLys 174
||| ||||| ||||| : : : : : ||| |||||
Qy 844 CAACACCTCAGGTGACCTCTCCAGTACGTGGATGACCTCTTCGCGCGGACCCACC 903
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 175 GlnAsnProAspIleValIleTyrGlnTyrMetAspLeuTyrValGly 191
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 904 AAACAGGACTGCTAGAGGACGAGGACACTACTGCTGGAATGCTGACCTAGGCTAC 963
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 192 -----SerAspLeuGluIle 196
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 964 AGAGCCTCTGTAAGAGCCCGAGATTTGCAGGAGAGGTAAACATCTTGGGTACAGT 1023
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 197 GlyGlnHisThrArgLysIleGluIleuArgGlnHisLeuLeuArgTyrGlyLeuThr 216
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1024 TTGGGGAGCGGCGACGCA 1053
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 217 ThrProAspLysLysHisGlnLysGluProPheLeuTyrMetGlyTyrGluLeuHis 236
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1054 GCACGGAAGAACTGTAGTCCAGATACCGGCCCAACC-----ACAGCCAAA 1101
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 237 ProAspLysTyrThrValGlnProIleValLeuProGluLysAspSerTyrThrValAsn 256
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1102 CABATGAGAGATTTTGGGACACCTGATTTTGCAGCTGTGGATCCCGGGT--- 1158
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 257 AspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSerGlnIleTyrProGlyIleLys 276
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1159 -----CGCAGCTTAGCAGCCCACTCTACCCGCTA 1188
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 277 ValArgGlnLeuCysLysLeuLeuArgGlyThrLysAlaLeuThrGluValIleProLeu 296
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1189 ACCAAA-----GAAAAGGGGAATTCCTCGGCTCTGAGCAC 1227
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 297 ThrGluGluAlaGluLeuAlaGluAsnArgGluIleLeuLysGluProValHis 316
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1228 CAGAAGGCAATTGATGCTATCAAAAGCCCTGTGAGCGCACCTGTCTGGCCCTCCCT 1287
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 317 GlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGlnGly 336
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1288 GACGTAACCTAACCTTACCTTTATGTGTGATGAGCGTAAAGGAGTAGCCCGGGAGTT 1347
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 337 GlnTrpThr-----TyrGlnIleTyrGlnLysAsnLeuLysThrGlyLys 354
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1348 TTAACCCAAACCTTAGGACCATGGAAGACCTGTGCTGCTCAAAAGAGCTCGAT 1407
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 355 TyrAlaArgMetArgGlyAlaHisThrAsnAspValLysGlnLeuThrGluAlaValGln 374
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1408 CCTGTAGCCAGTGTGGCCCATATGCTGAGGCTATCGCAGCTGTGGCCATCTGGTC 1467
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 375 LysIleThr-----GluSerIleValIleTrpGly 385
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1468 AAGGACCTGACAAATTGCTTTGGGACAGAAATATACTGTAATAGCCCCCATGCTG 1527
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 386 Lys-----ThrProLysPheLys 391
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1528 GAGAACATCGTTGGCAGCCCCCAGACCGATGGATGATGATGATGATGATGATGAT 1587
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 392 LeuProIleGlnLysGluThrTrpGluThrTrpTrpGlu-----TyrTrp 407
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1588 CAAAGCTGCTTCTCAGAGAGGGTCACTGCTCCACGAGCGGCTCTCAACCTGCC 1647
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 408 Gln-----Ala 409
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1648 ACTCTTCTGCTGAA-----GAGACTGATGAACGAGTCACTCATGATGATGATGAT 1698
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 410 ThrTrpIleProGluTrpGluPheValAsnThrProLeuValLysLeuTrpTyrGln 429
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1699 CTATGATGAGGAGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAA 1758
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 430 LeuGluLysGlu-----ProIleValGly--- 437
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1759 GTGCTAACCTGGTTCACTGACGAGGAGC----- 1788
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 438 AlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGly 457
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1789 TATCTGTGGAAAGGTAAAGAGATGGCTGGGCGCGGTGGTGGAGCGGACCGCACGATC 1848
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 458 TyrValThrAsnArgGlyArgGlnLysValValThrLeuThrAsp----- 472
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1849 TGGCCAGCAGCCTGCCGGAAGGAACTTCAGACACAAAGGCTGAGCTCATGCGCTCAGC 1908
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 473 -----ThrThrAsnGlnLysThrGluLeuGlnAlaIleTyr 484
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1909 CAAGCTTTGGCTGGCGGAGGGAATTCATAACATTTATACGACAGCAGGATGCTCC 1968
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 485 LeuAlaLeuGln---AspSerGlyLeuGluValAsnIleValThrAspSerGlnTyrAla 503
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 1969 TTTGCGACTGCACACGTACATGGGCGCATCTATAAACAAGGGGGTGTCTTACCTCACA 2028
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 504 LeuGly-----IleIleGlnAla 509
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2029 GGGAGGGAATAAGAAACAAGAGAAATTCCTAAGCCTATTAGAACGGCTGATACCTA 2088
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 510 GlnProAspGlnSerGluSerGluLeuValAsnGlnIleIleGluGlnLeuIleLysLys 529
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2089 AAAAGGCTAGCTATTATACACTGCTCTGGACATCAGAAAGCTAAGATCTCATATCCAGA 2148
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 530 GluLysValTyrLeuAlaTrpValProAlaHisLysGlyIle-----Gly 544
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 2149 GGAACACAGATGGCTGACCGGTGGCAAGCAGGCGAGCGGTGT 2196
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 545 GlyAsnGluGlnValAspLysLeuValSerAlaGlyIleArgLysVal 560
||| : : : : : ||||| : : : : : ||||| : : : : :
RESULT 8
US-11-022-562-221
; Sequence 221, Application US/11022562
; Publication No. US20050249742A1
; GENERAL INFORMATION:
; APPLICANT: Ruprecht, Ruth M.
; APPLICANT: Shisong, Jiang
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; FILE REFERENCE: A CVTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
; CURRENT APPLICATION NUMBER: US/11/022,562
; PRIOR FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: PCT/US03/20322
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392718
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 31, 97, 140, 141, 144, 178
; OTHER INFORMATION: Xaa = Any Amino Acid
US-11-022-562-221
Alignment Scores:
Pred. No.: 1,35e-09 Length: 210
Score: 217.50 Matches: 65
Percent Similarity: 47.9% Conservative: 37
Best Local Similarity: 30.5% Mismatches: 82
Query Match: 3.5% Indels: 29
DB: Gaps: 9
US-10-723-552-3_copy_2307_5741 (1-3435) x US-11-022-562-221 (1-210)
Qy 478 CAATCTCCCTCGAATACTCCCTGCTACCGGTTAGAAAGCGCTGGACTAATGACTATCGA 537
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 16 GluAsnProTyrAsnThrProValPheAlaIleLysLysLysAsp***ThrLysTrpArg 35
||| : : : : : ||||| : : : : : ||||| : : : : :
Qy 538 CCAGTACAGGACTTGAGAGAGGTCAATAAAGCGGTGCAGGAT-----ATACACCCA 588
||| : : : : : ||||| : : : : : ||||| : : : : :
Db 155 LysGlySerProAlaIlePheGlnSerMetThrLysIleLeuGluProPheArgLys 174
||| ||||| ||||| : : : : : ||| |||||
Qy 844 CAACACCTCAGGTGACCT

```
Db 36 LysLeuValAspPheArgGluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeu 55
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 589 ACAGTCCCGAACCTTATAACTCTTGTGTGCTCTCCACCCACCGAGAGCTGTATACA 648
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 56 GlyIleProHisProAlaGlyLeu-----LysLysLysSerValThr 70
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 649 GTATTGGACTTAAAGGATGCTCTTCTGCTGCTGAGATTACACCCACCTAGCACACCTT 708
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 71 ValLeuAspValGlyAspAlaTyrPheSerValProLeuAspGluSerPheArgLysTyr 90
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 709 TTTGCTTTTGAATGAGAGATCCAGGTACGGGAGAACCGGG--CAGCTCACCTGGAC 765
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 91 ThrAlaPheThrIlePro**ThrAsnAsnGluThrProGlyIleArgTyrGlnTyrAsn 110
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 766 CGACTGCCCGAAGGTTCAAGAACTCCCGGACCATCTTTGAGGAAGCCCTACACAGAC 825
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 111 ValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIle 130
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 826 CTGGCCAACTTCAGGATCCAAACCTCAGGTGACCTCTCCAGTACGTGGATGACCTG 885
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 131 LeuGluProPheArgIleLysAsnPro*****ValIle**GlnTyrMetAspAspLeu 150
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 886 -----CTTCTGCGGGAGCCACCAACAGGACTGCTTAGAAGGCCACGAAG 930
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 151 TyrValGlySerAspLeuGluIleGlyGlnHisArgAlaLysIleGluLeuArgLys 170
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 931 GCATCTACTGTGGAATGCTGACTAGGCTACAGAGCCTCTGTAAAGAGCCGAGATT 990
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 171 HisLeuLeu-----SerTrpGlyPhe**ThrProAspLysLysHisGln--- 185
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 991 TCCAGGAGAGGTAACATCTTG-----GGTACAGTTTGGGGAGCGGCGCATGG 1044
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 186 -----LysGluProProPheLeuTrpMetGlyTyrGluLeuHisPro---AspLysTrp 202
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 1045 CTGACGGAGCGGACCGAAGAAACTGTAGTCCAGATACCG 1083
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 203 ThrValGlnPro-----IleGlnLeuPro 210
    : : : : : | | | | | : : : : : | | | | | : : : : :
```

RESULT 9

```
US-11-230-251-26
; Sequence 26, Application US/11230251
; Publication No. US20060019322A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Caffkey, Robert
; APPLICANT: Ali, Shujath
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
; TITLE OF INVENTION: Genes
; FILE REFERENCE: DEX-0239
; CURRENT APPLICATION NUMBER: US/11/230,251
; CURRENT FILING DATE: 2005-09-19
; PRIOR APPLICATION NUMBER: US/09/957,708
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,746
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 26
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-230-251-26
```

Alignment Scores:

Pred. No.:	5,63e-09	Length:	422
Scores:	210.50	Matches:	88
Percent Similarity:	39.3%	Conservative:	52
Best Local Similarity:	24.7%	Mismatches:	109
Query Match:	3.4%	Indels:	107
DB:	7	Gaps:	18

```
US-10-723-552-3_copy_2307_5741 (1-3435) x US-11-230-251-26 (1-422)
QY 1982 CAAAAGCTGAGCTCATGGCCCTCACCAAGCTTTGGCGTGGCGAAGGAAATCCATA 1941
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 14 GlnArgAlaGluLeuValAlaValIleThrValLeuGln---AspPheAsnGlnSerIle 32
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 1942 AACATTTATACGGACAGCAGGTATGCTTT---CGGACTGCACACGTACATCGGCGCATC 1998
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 33 AsnIleValSerAspSerAlaTyrValValGlnAlaThrLysAspIleGluArgAlaLeu 52
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 1999 TATAAACAAAGGGGTTGCTTACCTCAGCAGGGAGGAAATAAAGAACAAAGAG----- 2052
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 53 IleLysTyr-----IleMetAspAspGlnLeuAsn 62
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2053 GAAATTTCTAAGCCTATTAGAGCCGTACATTTACCAAAAGG-----CTAGCTATT 2103
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 63 ProLeuPheAsnLeuLeuGlnGln---AsnValArgLysArgAsnPheProPheTyrIle 81
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2104 ATACACTGCTCTGGACATCAGAAAGCTAAAGACTCATATCCAGAGGAAACACAGATGGCT 2163
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 82 ThrHisIleArgAlaHisThrAsnLeuProGlyProLeuThrArgAlaAsnGluAla 101
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2164 GACCGGTTGCCAAGCAGGCGGCCCGGGTGTAACTTCTGCTTATATAAGAAATGCC 2223
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 102 AspLeuLeuValSerSerAla----- 108
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2224 AAAGCCCGAACCCAGACGACAGTACACCTAGAGACTGGCAAGAGATAAAAAAGATA 2283
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 108 -----PheMetGlu----- 108
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2284 GACCAGTTCTGAGACTCCGGAAGGACCTGTATACCTCAGATGGGAAGAAATCCTG 2343
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 109 -----PheMetGlu----- 111
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2344 CCCACAAAGAGGTTAGATATGTCCACAGATACATCTCTAACCCAC----- 2394
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 112 -----AlaGlnGluLeuHisAlaLeuThrHisValAsnAla 123
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2395 CTAGAACTAAACACCTGCAGCAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCTA 2454
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 124 IleGlyLeuLysAsnLysPheAspIleThrTrpLysGln----- 136
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2455 CCAGGAGTGGCTGACTCGGTGTCAAACATTTGTGCCCTGCCAGCTGTTTAAT----- 2508
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 137 -----ThrLysAsnIleValGlnHisCysThrGlnCysGlnIleLeuHisLeuAla 153
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2509 GCTAATCTTCCAGAAATGCTTCCAGGGAAGAGACTAAGGGGAAGCCACCCAGCGCTCAC 2568
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 154 ThrGlnGluAlaArgValAsnPro-----ArgGlyLeuCysProAsnValLeu 169
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2569 TGGGAAGTGGACTTCACTGAGGTAAGCGGCTAAATACGGGAACAAATACCTATTGCTT 2628
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 170 TrpGlnMetAspValMetHisVal---ProSerPheGlyLysLeuSerPheValHisVal 188
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2629 TTTGTAGACACTTTTTCAGGATGGGTAGAGCT---TATCCTACTTAAGAAAGACACTTCA 2685
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 189 ThrValAspThrTyrSerHisPheIleTrpAlaThrCysGlnThrGlyLysSerThrSer 208
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2686 ACCGTGTGGCTAAAAAATACTGGAAGAAATTTTCCAAGATTTGGAATACCTAAGGTA 2745
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 209 HisVal-----LysArgHisLeuLeuSerCysPheProValMetGlyValProGluLys 226
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2746 ATAGGTCAGACAATGGTCCAGCTTTTGTCCCGGTAAGTCAGGGACTGGCCCAAGATA 2805
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 227 ValLysThrAspAsnGlyProGlyTyr-----CysSerLysAlaValGlnLysPhe 243
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2806 TTGGGATTGATTGAAA-----CTGATTGTGCATACAGACCCCAAGCTCA 2853
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 244 LeuAsn---GlnTrpLysIleThrHisThrIleGlyIleLeuTyrAsnSerGlnGlyGln 262
    : : : : : | | | | | : : : : : | | | | | : : : : :
QY 2854 GGACAGGTAGAGAGATGAATAGAACCATTAAGAGACCCCTTACTATAA 2901
    : : : : : | | | | | : : : : : | | | | | : : : : :
Db 263 AlaIleIleGluArgThrAsnArgThrLeuLysAlaGlnLeuValLys 278
    : : : : : | | | | | : : : : : | | | | | : : : : :
```

RESULT 10

US-11-234-786-592
; Sequence 592, Application US/11234786
; Publication No. US20060024301A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.

TITLE OF INVENTION: PROSTATE-SPECIFIC POLYPEPTIDES AND FUSION

FILE REFERENCE: 210121.427C31

CURRENT APPLICATION NUMBER: US/11/234,786

PRIOR FILING DATE: 2005-09-23

PRIOR APPLICATION NUMBER: US 09/568,857

PRIOR FILING DATE: 2000-05-09

PRIOR APPLICATION NUMBER: US 09/536,857

PRIOR FILING DATE: 2000-05-27

PRIOR APPLICATION NUMBER: US 09/483,672

PRIOR FILING DATE: 2000-01-14

PRIOR APPLICATION NUMBER: US 09/439,313

PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: US 09/352,616

PRIOR FILING DATE: 1999-07-13

PRIOR APPLICATION NUMBER: US 09/288,946

PRIOR FILING DATE: 1999-04-09

PRIOR APPLICATION NUMBER: US 09/232,149

PRIOR FILING DATE: 1999-01-15

PRIOR APPLICATION NUMBER: US 09/159,812

PRIOR FILING DATE: 1998-09-23

PRIOR APPLICATION NUMBER: US 09/115,453

PRIOR FILING DATE: 1998-07-14

PRIOR APPLICATION NUMBER: US 09/030,607

PRIOR FILING DATE: 1998-02-25

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 701

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 592

LENGTH: 188

TYPE: PRT

ORGANISM: Homo sapien

US-11-234-786-592

Alignment Scores:

Pred. No.: 3,02e-06 Length: 188

Score: 174.50 Matches: 55

Percent Similarity: 47.1% Conservative: 33

Best Local Similarity: 29.4% Mismatches: 68

Query Match: 2.8% Indels: 31

DB: 7 Gaps: 10

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-11-234-786-592 (1-188)

QY 2371 CAACAGATACATCGTCTTAACCCAC-----CTAGGAACATAAACACCTGCAGCAGTTG 2421

Db 19 GlnGluLeuHisAlaLeuThrHisValAsnAlaIleGlyLeuLysAsnLysPheAspIle 38

QY 2422 GTCAGAACATCCCTTATCATGTCTGAGCTACAGAGTGGCTGACTGCTGGTGGTCAAA 2481

Db 39 ThrTrpLysGln-----ThrLysAsnIleValGln 48

QY 2482 CATTGTGTCCTGCCAGCTGGTTAAAT-----GCTAATCCTCTTCAGAAATGCTCCAGGG 2535

Db 49 HisCysThrGlnCysGlnIleLeuHisLeuAlaThrGlnGluAlaArgValAsnPro--- 67

QY 2536 AAGAGACTAAGGGAAGCCACCCAGCGCTCACTGGGAAGTGGACTTCACCTAGAGTAAAG 2595

Db 68 -----ArgGlyLeuCysProAsnValLeuTrpGlnMetAspValMetHisVal--- 83

QY 2596 CCGGCTAAATACGGAACAAATACCTATTGGTGTGTTTGTAGACACCTTTTCAGGATGGGTA 2655

Db 84 ProSerPheGlyLysLeuSerPheValHisValThrValAspThrTyrSerHisPheIle 103

QY 2656 GAGGCT---TATCCTACTAAGAAAGAGACTTCAACCGTGGTGGCTAAAAAAATACTGGAA 2712

Db 104 TrpAlaThrCysGlnThrGlyGluSerThrSerHisVal-----LysArgHisLeuLeu 121

QY 2713 GAAATTTTCCAGATTTGGAATACCTAAGTAATAGGTGAGGTGAGCAATGGTCCAGCTTTT 2772

Db 122 SerCysPheProValMetGlyValProGluLysValLysThrAspAsnGlyProGlyTyr 141

QY 2773 GTTCCCGAGGTAAAGTCAAGGACTGGCCCAAGATATTGGGGATTGGTAA----- 2823

Db 142 CysSerLysAlaPheGln-----LysPheLeuAsn---GlnTrpLysIleThrHis 157

QY 2824 ---CTGCATTGTGCATACAGACACCCAAAGCTCAGGACAGGTAGAGAGGATGAATAGAAC 2880

Db 158 ThrIleGlyIleLeuTyrAsnSerGlnGlyGlnAlaIleIleGluGlyThrAsnArgThr 177

QY 2881 ATTAAGAGACCTTTACTAAA 2901

Db 178 LeuLysAlaGlnLeuValLys 184

RESULT 11

US-11-175-689-9

; Sequence 9, Application US/11175689

; Publication No. US20060024806A1

; GENERAL INFORMATION:

; APPLICANT: STINEAR, TIMOTHY P.

; APPLICANT: COLE, STEWART T.

; APPLICANT: LEADLAY, PETER F.

; APPLICANT: SMALL, PAMELA L.C.

; APPLICANT: JOHNSON, PAUL D.R.

; APPLICANT: JENKIN, GRANT A.

; APPLICANT: DAVIES, JOHN K.

; APPLICANT: HAYDOCK, STEPHEN F.

; TITLE OF INVENTION: THE MYCOLACTONE LOCUS: AN ASSEMBLY LINE FOR PRODUCING

; FILE REFERENCE: NOVEL POLYKETIDES, THERAPEUTIC AND PROPHYLACTIC USES

; FILE REFERENCE: 03495.0329-01

; CURRENT APPLICATION NUMBER: US/11/175,689

; CURRENT FILING DATE: 2005-07-07

; PRIOR APPLICATION NUMBER: 10/987,592

; PRIOR FILING DATE: 2004-11-15

; PRIOR APPLICATION NUMBER: 60/519,864

; PRIOR FILING DATE: 2003-11-14

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 3.3

; SEQ ID NO 9

; LENGTH: 14130

; TYPE: PRT

; ORGANISM: Mycobacterium ulcerans

US-11-175-689-9

Alignment Scores:

Pred. No.: 0.00134 Length: 14130

Score: 146.50 Matches: 243

Percent Similarity: 31.0% Conservative: 148

Best Local Similarity: 19.2% Mismatches: 399

Query Match: 2.4% Indels: 473

DB: 7 Gaps: 70

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-11-175-689-9 (1-14130)

Db 7996 SerArgIleHisThrLeuThrArgGlnThrLeuThrValValGlnAspTrpLeuThrHis 8015
Qy 1957 -----AGCAGGTATGCTTTCGCACTGCACAC----- 1983
Db 8016 ProAspThrThrGlyThrArgLeuValIleValThrArgHisGlyValSerThrSerAla 8035
Qy 1984 -----GTACATGGGCGCATCTATAAACAAGGGGTTCCTTACC 2022
Db 8036 HisAspProValProAspLeuAlaHisAlaValTrp-----GlyLeuIleArg 8052
Qy 2023 TCAGCAGGAGGAAATAAGAACAAACAGAGGAATCTA----- 2061
Db 8053 SerAlaGlnAsnGluHisProGlyArgPheThrLeuLeuAspThrAspAsnThrAsn 8072
Qy 2062 -----AGCCTATTAGAACCGGTACATTTACCAAAAGG-----CTAGCTATT--- 2103
Db 8073 SerAspThrLeuThrAlaLeuThrLeuProThrArgGluAsnGlnLeuAlaIleArg 8092
Qy 2104 -----ATACACTGCTCGGACATCAGAAAGCTAAAGATCTCATATCCAGAGGAAC 2154
Db 8093 ArgAspThrIleHisIlePro----- 8099
Qy 2155 CAGATGGCTGACCGGTTGCCAAGCAGCAGCCAGGGGTGTTAACTTCTGCTTATAATA 2214
Db 8100 -----ArgLeuThrArgHisSerSerAspGlyAlaLeuThrAlaProValVal 8115
Qy 2215 GAAATGCCAAAGCCCCAGAACCCAGACAGTACACCCCTAGAAGACTGGCAAGAGATA 2274
Db 8115 ----- 8115
Qy 2275 AAAAGATAGACCAAGTCTCTGAGACTCCGGAAGGAGCTGCTATACCTCAGATGGGAAG 2334
Db 8116 -----ValAsp-----ProGluGlyThrValLeuIleThrGlyGlyThr 8128
Qy 2335 GAAATCTGCCCCACAAAGAGGGTTAGATATATGCCAAGATACATCGTCTCAACCCAC 2394
Db 8129 GlyThrLeu-----GlyAlaLeuPheAlaGlu-----HisLeuValSerAla 8142
Qy 2395 CTAGGAATAAACACCTGCAGCAGTGTGTGAGAACATCCCTTATCATGTTCTCAGAGCTA 2454
Db 8143 HisGlyValArgHisLeuLeuLeuThrSerArgArgGlyPro-----GlnAla 8158
Qy 2455 CCAGAGTGGCTGATCGGTGGTCAAAACATGTGTGCTGCCCTGCCAGCTGGTTAATGCTAAT 2514
Db 8159 HisGlyAlaThrAsp----- 8163
Qy 2515 CCTTCCAGATGCTCCAGGGAAGAGACTAAGGGGAAGCCACCAGCGGCTCACTGGGAA 2574
Db 8164 -----LeuGlnGlnArgLeuThrAspLeuGlyAlaHisValThr 8176
Qy 2575 GTGACTTCTACTGAGTAAAGCCGGCTAAATACGGAACAAATACCTATTGTTGTTTGTGA 2634
Db 8177 IleThrAlaCysAspIleSerAspPro----- 8185
Qy 2635 GACACCTTTTCAGGATGGTAGAGGCTTATCCTTACTAAGAAAGAGACTTCAACCGTGGTG 2694
Db 8186 GluAlaLeuAlaLeuValAsnSerValProThrGlnHisArgLeuThrAlaValVal 8205
Qy 2695 GCTAAAAAATCTCGGAAGAAATTTTCCAGATTTTGGATACCTAGGTAATAGGTCA 2754
Db 8206 HisThrAlaAlaValLeuAlaAspThrPro-----ValThrGluLeuThrGlyAsp 8222
Qy 2755 GACAATGGTCCAGCTTTTGTGTCAGGTAAGTCAAGGACTGGCCCAAGATATTGGGGATT 2814
Db 8223 -----GlnLeuAspGlnValLeuAlaProLysIleAspAla 8234
Qy 2815 GATTGGAACTGCATTGTGCATACAGACCCCAAGAGCTCAGACAGCTAGAGAGATGAAT 2874
Db 8235 AlaTrpGlnLeuHis----- 8239
Qy 2875 AGAACCTTAAGAGACCCCTTACTAATTTGACCGGAGAGACTGGCGTTAATGATGGATA 2934
Db 8240 -----GlnLeuThrTyrGluHisAsn----- 8246

Qy 2935 GCTCTCTGCTCTTTTGTGCTTTTATAG-----GTTAGGAACACCCCTGGACAG 2982
Db 8247 ---LeuSerAlaPheIleMetPheSerSerMetAlaGlyMetIleGlySerProGlyGln 8265
Qy 2983 -----TTT 2985
Db 8266 GlyAsnTyrAlaAlaAlaAsnThrAlaLeuAspAlaLeuAlaAspTyrArgHisArgLeu 8285
Qy 2986 GGGCTGACCCCTATGAATTACTCTACGG-----GGACCCCCCCCA 3027
Db 8286 GlyLeuProAlaThrSerLeuAlaTrpGlyTyrTrpGlnThrHisThrGly----- 8302
Qy 3028 TTGCTAGAAATGCTCTGTACATAGTGTGACGTGCTTCCAGCCCTTGTCTCTCT 3087
Db 8303 -----LeuThrAlaHisLeuThrAspValAspLeuAlaArg-----MetThr 8316
Qy 3088 AGGCTCAAGCAGCTTGTAGTGGTGAGACAACAGAGCGTGGAGCAACTCCGGAGGCTTAC 3147
Db 8317 ArgLeuGlyLeuMetProIleAlaThrSerHisGlyLeuAlaLeuPheAspAlaAlaLeu 8336
Qy 3148 TCAGGAGGAGAGACTTGCAGATCCCATCGTTTCCAGTGGGAGATTTCAGTCTACGTT 3207
Db 8337 AlaThrGlyGlnProValSerIlePro-----AlaProIleAsnThrHisThrLeuAla 8354
Qy 3208 AGACGCCACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGCGCTTATCTC---GTACTT 3264
Db 8355 ArgHisAlaArgAspAsnThrLeu-----AlaProIleLeuSerAlaLeu 8369
Qy 3265 TTGACCACACCA-----ACGGCTGTGAAAGTCGAA 3294
Db 8370 IleThrThrProArgArgAlaAlaSerAlaAlaThrAspLeuAlaAlaArgLeuAsn 8389
Qy 3295 GGAATCTCC 3303
Db 8390 GlyLeuSer 8392

RESULT 12
US-11-175-689-7
; Sequence 7, Application US/11175689
; Publication No. US20080024806A1
; GENERAL INFORMATION:
; APPLICANT: STINEAR, TIMOTHY P.
; APPLICANT: COLE, STEWART T.
; APPLICANT: LEADLAY, PETER F.
; APPLICANT: SMALL, PAMELA L.C.
; APPLICANT: JOHNSON, PAUL D.R.
; APPLICANT: JENKIN, GRANT A.
; APPLICANT: DAVIES, JOHN K.
; APPLICANT: HAYDOCK, STEPHEN F.
; TITLE OF INVENTION: THE MYCOLACTONE LOCUS: AN ASSEMBLY LINE FOR PRODUCING
; TITLE OF INVENTION: NOVEL POLYKETIDES, THERAPEUTIC AND PROPHYLACTIC USES
; FILE REFERENCE: 03495.0329-01
; CURRENT APPLICATION NUMBER: US/11/175,689
; PRIOR FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: 10/987,592
; PRIOR FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: 60/519,864
; PRIOR FILING DATE: 2003-11-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 7
; LENGTH: 16990
; TYPE: PRT
; ORGANISM: Mycobacterium ulcerans
US-11-175-689-7

Alignment Scores:
Pred. No.: 0.00241 Length: 16990
Score: 143.50 Matches: 242
Percent Similarity: 31.4% Conservative: 151
Best Local Similarity: 19.3% Mismatches: 409
Query Match: 2.3% Indels: 451

```
DB: 7 Gaps: 69
US-10-723-552-3_copy_2307_5741 (1-3435) x US-11-175-689-7 (1-16990)
Qy 40 CGAAGAACAGTTGACTTGGGAGTGGGACGGGTAAACCCACTCGTCTTCTGGTCATCACTGAG 99
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15848 ArgArgThrThrArgLeuGlnValSerHisAlaPheHisSer-----ProHis 15863
Qy 100 TGCCAGACACCCCTCTTAGGTAGAGACTTATTGACCAAGATGGGAGCACAATTTCTTTT 159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15864 MetAsp-----ProIleLeu-----GluGlnPheArgGlnIleAlaAlaGlnLeuThrPhe 15880
Qy 160 GAACAAGGGAACACAGAGTGTCTGCAAAATACAAACCTATC--ACTGTGTGACCCCTC 216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15881 -----SerAlaProThrLeuProIleLeuSerAsnLeuThrGly 15893
Qy 217 CAATTAGATGACGAATATCGACTACTCTCCCTCTAGTAAAGCCGTCAAAATATACAA 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15894 GlnIleAlaArgHisAspGlnLeuAlaSerPro-----Asp 15905
Qy 277 TTCTGGTTGGACAG-----TTTCCCAAGCCTGGGCGAGAACCGCA 318
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15906 TyrTrpThrGlnGlnLeuArgAsnThrValArgPheHisAspThrValAlaAlaLeuLeu 15925
Qy 319 GGGATGGCTTTGGCAAGCAAGTTCCTCCCAACAAGTTATTCAACTGAAGCCAGTGCCACA 378
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15926 GlyAlaGly-----GluGlnValPheLeuGluLeuSerProHis 15938
Qy 379 CCAGTGTCACTCAGACAGTACCCCTTGAGTAAAGAGCTCAAGAGGAATTCGGCCGCAT 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15939 ProVal-----LeuThrGlnAlaIleThrAspThr 15948
Qy 439 GTCCAAAGATTATCCACAGGGCATCTAGTTCTCT----- 474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15949 ValGluGlnAlaGlyGlyGlyAlaAlaValProAlaLeuArgLysArgProAsp 15968
Qy 475 -----GTCCAATCTCCCTGG 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15969 AlaValAlaPheAlaAlaLeuGlyGlnLeuHisCysHisGlyIleSerProSerTrp 15988
Qy 490 AATACT-----CCCCTG--CTACCGGTTAGAAAG----- 516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15989 AsnValLeuTyrCysGlnAlaArgProLeuThrLeuProThrTyrAlaPheGlnHisGln 16008
Qy 517 -----CCTGGGACTAATGACTATCGACCATCGACAGTACAGGACTTCAGAGAGTGC 561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16009 ArgTyrTrpLeuLeuProThrAlaGlyAspPheSer----- 16020
Qy 562 AATAAAGGGTCAGGATATACACCAACAGTCCCGAACCCCTTATAAACCCTCTTGTGCT 621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16021 GlyAlaAsnThrHisAlaWechHisProLeuLeu-----AspThrAlaThrGlu 16036
Qy 622 CTCGCCACCCCAACGGAGCTGTATACAGTATTGGAATTTAAGGATGCGCTTCTTCTGCGCTG 681
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16037 LeuAlaGluAsnArgGlyTrp-----ValPheThrGly 16047
Qy 682 AGATTACACCCACTAGCCACCACTTTTTCGCTTCGAATGGAGAGATCCAGGTACGGGA 741
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16048 ArgIleSerProArgThrGlnPro-----TrpLeuAsnGluHisAlaVal 16062
Qy 742 AGAACGGGCAGCTCACCTGGACCGGACTGCCCAAGGGTTCAAGAACTCCCGACCATC 801
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16063 GluSerAlaValLeu-----PheProAsnThrGlyPheVal 16074
Qy 802 TTTCGAGAGCCCTACACAGAGACCTGGCCAACTTCAGGATCCCAACCCCTCAGGTGACC 861
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16075 ---GluLeuAlaLeuHis-----ValAlaAsp---ArgAlaGlyTyrSerSerValAsn 16090
Qy 862 CTCCTCCAGTAGCTGGATGACTGCTTCTGGCGGAGGCCCAACACAGGACTGCTTAGAA 921
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16091 GluLeuIleValHisThrProLeuLeuLeuAlaGlyHisAspThrAlaAsp----- 16107
Qy 922 GGCAGGAAGGCCTACTCTGTGAAATTGCTGACCTAGGCTACAGAGCCTCTGCTAAGAAG 981
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
Db 16408 AlaHisGlnAlaSerAspThrAlaValSerSerArgIleHisThrLeuThrArgGlnThr 16427
Qy 1939 ATAAACATTATACCGAC-----AGCAGGTATGCCCTT 1971
Db 16428 LeuThrValGlnAspThrLeuThrHisProAspThrThrGlyThrArgLeuValIle 16447
Qy 1972 GCGACTGCACAC-----GTACATGGG 1992
Db 16448 ValThrArgHisGlyValSerThrSerAlaHisAspProValProAspLeuAlaHisAla 16467
Qy 1993 GCCATCTATAACAAGGGTGTCTTACCTCAGCAGGAGGGAATAAAGCAACAAGAG 2052
Db 16468 AlaValTrp-----GlyLeuIleArgSerAlaGlnAsnGluHisProGlyArgPhe 16484
Qy 2053 GAAATTTCTA-----AGCCTATTAGAACCGGTACATTTA 2085
Db 16485 ThrLeuLeuAspThrAspAsnThrAsnSerAspThrLeuThrThrAlaLeuThrLeu 16504
Qy 2086 CCAAAAAGG-----CTAGCTATT-----ATACACTGTCCTGGACATCAG 2124
Db 16505 ProThrArgGluAsnGlnLeuAlaIleArgArgAspThrIleHisIlePro----- 16521
Qy 2125 AAGCTAAAGATCTCATATCCAGAGAAACCCAGATGGCTGACCGGTTGCCAAGCAGCA 2184
Db 16522 -----ArgLeuThrArgHisSer 16527
Qy 2185 GCCCAGGTGTAACTTCTGCCTATAATAGAAATGCCAAAGCCCAAGACCCAGACGA 2244
Db 16528 SerAspGlyAlaLeuThrAlaProValVal----- 16537
Qy 2245 CAGTACACCTTAGAAGACTGCCAAGAGATAAAAGATAGACCAAGTCTCTGAGACTCCG 2304
Db 16538 -----ValAsp-----Pro 16540
Qy 2305 GAAGGACCTGTATACCTCAGATGGGAGGAATCTCGCCCAACAAAGAGGGTTAGAA 2364
Db 16541 GluGlyThrValLeuIleThrGlyGlyThrGlyThrLeu-----GlyAlaLeu 16556
Qy 2365 TATGTCACACAGATACATCGTCTAACCCACCTAGAACATAACACCTGCAGCAGCTGGTC 2424
Db 16557 PheAlaGlu-----HisLeuValSerAlaHisGlyValArgHisLeuLeuThrSer 16574
Qy 2425 AGAACATCCCTTATCATGTTCTCAGGCTACACAGAGTGGCTGACTCGGTGGTCAAAAT 2484
Db 16575 ArgArgGlyPro-----GlnAlaHisGlyAlaThrAsp----- 16585
Qy 2485 TGCTGCGCTGCCAGCTGTTAATGCTTAATCTTCCAGATGCCTCCAGGAAGACACTA 2544
Db 16586 -----LeuGlnGln 16588
Qy 2545 AGGGGAAGCCACCCAGCGCTCACTGGGAAGTGGACTTCACTAGAGTAAAGCCGCTAAA 2604
Db 16589 ArgLeuThrAspLeuGlyAlaHisValThrIleThrAlaCysAspIleSerAspPro--- 16607
Qy 2605 TACGGAAACAAATACCTATTGGTTTGTGTAGACACCTTTTCAGGATGGGTAGAGGCTTAT 2664
Db 16608 -----GluAlaLeuAlaLeuValAsnSerVal 16617
Qy 2665 CCTACTAAGAAGAGACTTCAACCGTGGTGGCTAAAAAATACTGGGAAGAAATTTTCCA 2724
Db 16618 ProThrGlnHisArgLeuThrAlaValValHisThrAlaAlaValLeuAlaAspThrPro 16637
Qy 2725 AGATTTGGAATACCTAAGGTAAATAGGTGCAGACAATGGTCCAGCTTTTGTGGCCAGGTA 2784
Db 16638 -----ValThrGluLeuThrGlyAsp-----GlnLeu 16646
Qy 2785 AGTCAGGAGCTGCCCAAGATATTGGGATGATTGGAACACTGCATTTGTGCATACAGACCC 2844
Db 16647 AspGlnValLeuAlaProLysIleAspAlaAlaIleArgGlnLeuHis----- 16661
Qy 2845 CAAAGCTCAGGACAGGTAGAGAGGATGAATGAACCACTTAAGAGACCCCTTACTAAATTG 2904
Db 16662 -----GlnLeu 16663

Qy 2905 ACCCGGAGACTGGGTAAATGATTGATAGCTCTCTCTGCCCTTTGTCTTTTAGG--- 2961
Db 16664 ThrTyrGluHisAsn-----LeuSerAlaPheIleMetPheSerSer 16677
Qy 2962 -----GTTAGAACACCCCTGGACAG----- 2982
Db 16678 MetAlaGlyMetIleGlySerProGlyGlnGlyAsnTyrAlaAlaAsnThrAlaLeu 16697
Qy 2983 -----TTTGGGTGACCCCTATGAATTACTCTACCGG 3015
Db 16698 AspAlaLeuAlaAspTyrArgHisArgLeuGlyLeuProAlaThrSerLeuAlaIleProGly 16717
Qy 3016 -----GGACCCCCCATTTGGTAGAAATTTGTTCTGTACATAGTCTCT 3057
Db 16718 TyrTrpGlnThrHisThrGly-----LeuThrAlaHisLeuThr 16730
Qy 3058 GACGTGCTGCTTTCCACGCTTTTCTCTAGGCTCAAGGACCTTGAGTGGGTGAGACAA 3117
Db 16731 AspValAspLeuAlaArg-----MetThrArgLeuGlyLeuMetProIleAlaThrSer 16748
Qy 3118 CGACGTGGAGGCAACTCCGGGAGGCTTACTCAGGAGGAGGAGACTTGCAGATCCCAT 3177
Db 16749 HisGlyLeuAlaLeuPheAspAlaAlaLeuAlaThrGlyGlnProValSerIlePro--- 16767
Qy 3178 CGTTTCCAGTGGGAGATTCTAGTCTAGTACGCCACCGTCAGGAAACCTCGAGACT 3237
Db 16768 ---AlaProIleAsnThrHisThrLeuAlaArgHisAlaArgAspAsnThrLeu----- 16784
Qy 3238 CGGTGGAGGCGCTTATCTC---GTACTTTTGACACACCA----- 3276
Db 16785 -----AlaProIleLeuSerAlaLeuIleThrProArgArgAlaAlaSer 16801
Qy 3277 -----ACGGCTGTGAAAGTCGAAGGATCTCC 3303
Db 16802 AlaAlaThrAspLeuAlaAlaArgLeuAsnGlyLeuSer 16814

RESULT 13
US-11-143-980-47
; Sequence 47, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: Hucul, John
; APPLICANT: Haltli, Bradley A.
; APPLICANT: Wagenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; TITLE OF INVENTION: Polyketide
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; PRIOR FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 47
; LENGTH: 5712
; TYPE: PRF
; ORGANISM: Streptomyces sp.
US-11-143-980-47

Alignment Scores:
Pred. No.: 0.0394 Length: 5712
Score: 126.50 Matches: 259
Percent Similarity: 29.8% Conservative: 150
Best Local Similarity: 18.9% Mismatches: 455
Query Match: 2.0% Indels: 511
```

DB: 7 Gaps: 70
US-10-723-552-3_COPY_2307_5741 (1-3435) x US-11-143-980-47 (1-5712)
QY 40 CGAAGACAGTT-----GACTGGGAGTGGACGG 69
DB ArgArgThrValArgPheLeuAspGlyValArgCysLeuGluAspGluGlyValThrThr 849
QY 70 GTAACCCACTGTTCTGGTGTATACCTGAGTGGCCAGCACCCCTCTTAGGTAGAGACTTA 129
DB 11leuGlu-----LeuGlyProAspGlyAlaLeuThrThrLeuAlaAspCys 866
QY 130 TTGACCAAGATGGGAGCAAAATT---TCTTTGAACAGGGAACACAGAGTGTCTGCA 186
DB LeuThrGlyProGlyThrLeuValGlyThrLeuArgArgAspArgProGluProGlnAla 886
QY 187 AATAACAAACCTATCACTGTGTGACC---CTCCAAATTAGATGACGAATATCGACTATAC 243
DB 887 -----LeuValThrAlaLeuAlaGluLeuTyrValSerGlyValGluValAlaTyr 903
QY 244 TCTCCCTAGTAAGCCTGATCAAAATATA----- 273
DB SerProLeuValSerGlyValArgGileProLeuProThrTyrAlaPheGlnArgGln 923
QY 274 CAATTCTGTTGGACAGTTTCCCAAGCCTGGGCAGAAACCGCAGGATGGTGTGGCA 333
DB ArgTyrTrpPheSerAlaProGlyProGluSerGlyThrThrProGlyHisGlyValThr 943
QY 334 AAGCAAGTTCGCCCAAGTTATTCAACTGAAGGCCAGTGGCCACACCACTGTCTAGTCTAGA 393
DB SerGlyArgGluArgThrAspThrGlyLeuSerGlyAspGluAlaProAspThr----- 961
QY 394 CAGTACCCCTGAGTAAGACAGCTCAAGAGGAATTCGGCGCATGTCTCAAGATTATATC 453
DB 962 ---GlyProSerGlyGlyGluThrLeuGlyMetValArgAlaHisAlaAlaValValLeu 980
QY 454 CAACAGGGCATCTAGTTCCTGCTCCCAATCTCCCTGGNAAT----- 492
DB 981 GlyTyrAlaSerAlaThrAlaAlaGlyAlaGluHisThrPheLeysGlnLeuGlyPheAsp 1000
QY 493 -----ACTCCCTCTCTACCGGTTAGA 513
DB 1001 SerIleThrAlaValGluLeuCysGluArgLeuGlyAlaAlaThrAlaLeuProLeu--- 1019
QY 514 AAGCCTGGGACT-----AATGACTAT---CGACCACTACAGGACTTCAGAGAGTCAAT 564
DB 1020 ---ProGlyThrLeuLeuPheAspTyrProThrProAlaAlaLeuAlaGluHisLeuHis 1038
QY 565 AAACGGGTG-----CAGGATATACACCCCAACAGTCCCGAACCT--- 603
DB 1039 ArgArgLeuHisGlyArgThrAspGluGlnAlaAlaProAlaThrValProThrProAsp 1058
QY 604 -----TATAACCTCTTGTGTCTCTCCCA----- 627
DB 1059 GlyGlyAspProValValIleValGlyMetGlyCysArgPheProGlyArgAlaHisSer 1078
QY 628 CCCCAACGGAGCTGTATACAGTATTGCACTTAAGGATGCTTC----- 672
DB 1079 ProGluAspLeuTrpArgIleValAlaAspGlyGluAspAlaIleSerGlyPheProSer 1098
QY 673 -----TTCTGCTGAGATTACACCCCACTAGCCACCACTTTTTCCTTC 717
DB 1099 AspArgGlyTrpAspLeuAlaGlyLeuTyrHisProAspProAsp----- 1113
QY 718 GAATGGAGAGATCCAGGTACG-----GGAGAACCGGGCAGCTCACCTGGACCCGACTG 771
DB 1114 -----HisProGlyThrSerTyrAlaArgAspGly----- 1123
QY 772 CCCCAAGGGTTCAAGAACTCCCGGACCATCTTTGACGAAGCCCTACACAGAGACCTGGCC 831
DB 1124 -----GlyPhe-----LeuTyrAsp-----AlaAlaGluPheAspAlaGly 1135
QY 832 AACTTCAGGATCCAAACACCTCAGGTGACCCCTCTCCAGTACGTAGTGGATGACCTGCTTCTG 891

DB 1136 PhePheGlyIleSerProArgGluAlaGluAlaMetAspProGlnGlnArgLeuLeu 1155
QY 892 -----GCGGGAGCCACCAACAGAGGACTGTTTGAAGGC 924
DB GluThrSerTrpGluAlaLeuGluArgAlaGlyIleProAlaGluHisIleLeysGlySer 1175
QY 925 ACRAAGGCACCTACTGCTGGAATGCTGACTAGCTACAGAGCTCTGCTAAAGAAG--- 981
DB 1176 SerThrGlyValPheIleGlyAlaSerSerValGlyTyrAlaAlaAspAlaGlyGluAla 1195
QY 982 -----GCCCAGATTTCAGGAGAGAGGTAACTATAC 1011
DB 1196 AlaGluGlyTyrGlnLeuThrGlyThrAlaAlaSerValAlaSerGlyArgValSerTyr 1215
QY 1012 ---TTGGGG----- 1017
DB ThrLeuGlyLeuGluGlyProAlaValThrValAspThrAlaCysSerSerSerLeuVal 1235
QY 1018 -----TACAGTTTCGGGACGGCAGCGATGGCTGACGGAGGCACGG 1059
DB AlaLeuHisLeuAlaValGlnSerLeuArgAlaGlyGlnCysSerLeuAlaLeuAlaGly 1255
QY 1060 AAGAAAACTGTAGTCCAGATACCGGCCCAACACAGCCAAACAAATGAGAGAGTTTTG 1119
DB GlyValThrValMetAlaThrProAlaMetPheValGluPheSerArgGlnArgGlyLeu 1275
QY 1120 GGGACAGCTGGATTTTCAGACTG-----TGGATCCCG 1152
DB AlaMetAspGlyArgCysLysAlaPheAlaAlaAlaAlaAspGlyThrGlyTrpAlaGlu 1295
QY 1153 GGGTTTCGACTTAGCAGCCCACTCTACCCGCTAACCAAGAAAAAGGGGAATTTCTCC 1212
DB GlyValGlyValLeuValVal----- 1302
QY 1213 TGGGCTCCTGAGCACCAGAGCGCATTTGATGCTATCAAAAAG-----GCC 1257
DB 1303 -----GluArgLeuSerAspAlaGluArgAsnGlyHisArgValLeuAla 1317
QY 1258 CTGCTGAGCGCACCTGCTCTGCGCCCTCCTGACGTAAACCTTTTACCCTTTATGTG 1317
DB 1318 ValValArgGlySerAlaVal----- 1324
QY 1318 GATGAGCGTAAGCGAGTAGCCGGGGAGTTTAAACCAACCCCTAGGACCATGGAGAAGA 1377
DB 1325 ---AsnGlnAspGlyAlaSerAsnGly---LeuThrAlaProAsnGlyPro----- 1339
QY 1378 CCTGTGCGCTACCTGTCTCAAGAGAGCTGATCCTCTGTAGCCAGTGTGGCCCATATGCTG 1437
DB 1340 -----SerGlnArgValIleArg 1346
QY 1438 AAGCTATTCGACGTGTGGCCATATCTGTCAGGACCGCTGACAAATTG----- 1485
DB GlnAlaLeuAlaSerAlaGlyLeuValAlaSerAspValAspAlaValGluAlaHisGly 1366
QY 1486 -----ACTTTGGGACAGAAATATA-----ACTGTAATAGCCCCCATGTCATTG 1527
DB ThrGlyThrThrLeuGlyAspProIleGluAlaGlnAlaLeuLeuAlaThrTyrGlyGln 1386
QY 1528 GAGAACATCGTTTCGGCAGCCCCCAGACCGATGGATG-----ACCAACCGCCGATGACC 1581
DB GlyArgAspAlaAspArgPro-----LeuTrpLeuGlySerValLysSerAsnIleGly 1404
QY 1582 CACTATCAAGC-----CTGTTCTCACAGAGGGTGCAGTTCGCTCCACCAGCC 1632
DB HisThrGlnAlaAlaAlaGlyValIleLysMetValMetAlaMetArgHis 1424
QY 1633 GCTCTCAACCTGCCACTCTTCTGCTGAGAGAGACTGATGAACCCAGTGCATCATGATTGC 1692
DB GlyValLeuProArgThrLeu-----HisValAspGluProSerThrHis----- 1439
QY 1693 CATCACTATTGATTGAGGAGACTGGGTCCGCAAGGACCTTTACAGACATACCGCTGACT 1752

Db 1440 -----ValAspTrpSerGlyGlyArgValGluLeu-----LeuThr 1451
Qy 1753 GGAGAGTGTCTAACCTGTTCACTGACGGAAGCAGCTATGTGTGGTGAAGGTAAGAGGATG 1812
Db 1452 Gly---ThrThrProTrpProThrGly-----GlyLeuArgArg 1464
Qy 1813 GCTGGGGGGGGTGTG-----GTGACGGGACCCGACG-----ATCTGGGCCAGCAGGCTG 1863
Db 1465 AlaGlyValSerSerPheGlyValSerGlyThrAsnAlaHisValLeuGluGlnVal 1484
Qy 1864 CCGGAA-----GGAACTTCACGACCAAAAG 1887
Db 1485 ProGluThrAlaArgProThrGlyProIleGlyGluAspArgGlyGluAlaAlaProVal 1504
Qy 1888 GCTGAGCTCATGGCC-----CTCAGCGAAGCTTTCGGCTGGCC 1926
Db 1505 AlaTrpValLeuSerGlyGlnGlyGluThrGlyLeuArgAlaGlnAlaGluArgLeuCys 1524
Qy 1927 GAAGGGAATCCATAAATTTAT-----ACGACAGCAGGTATGCTTTGGGACTGACAC 1983
Db 1525 AlaPheMetAlaAlaAspThrArgProThrProAlaGluValGlyTrpSerLeuAlaSer 1544
Qy 1984 GTACATGGGGCCATCTATAACAAGGGGTGTCTTACCTCAGCAGGAGGGAATAAAG 2043
Db 1545 ThrArg---AlaThrLeuSerHisArgAlaValValGlyAlaGlyArg----- 1560
Qy 2044 AACAAAGAGGAATTTCTAAGCCTATTAGAAGCCGTACATTTTACCAGAAAGGCTAGCTATT 2103
Db 1561 -----AspGluLeuLeuArgGlyValAsnAlaVal----- 1570
Qy 2104 ATACACTGTCTCGACATCAGAAAGCTAAAGATCTCATATCCAGAGAAACCATAGTGCT 2163
Db 1570 ----- 1570
Qy 2164 GACCGGTTCAGAGCAGCCAGCGGTGTAACTTCTGCTCTATAATAGAAATGCC 2223
Db 1571 -----AlaAsnGlyThrProValProGlyValVal----- 1580
Qy 2224 AAAGCCCCAGAACCCAGACAGTACACCCCTAGAGACTGGCAAGAGATAAAAAAGATA 2283
Db 1580 ----- 1580
Qy 2284 GACCAGTTCTGTAGACTCCGGAAGGAGCCTGTATACCTCAGATGGAAGGAATCCTG 2343
Db 1581 -----ArgGlyThrGlyAlaSerGlyAspValValPheValPhe 1593
Qy 2344 CCCACAAAGAGGTAGATATGTCCACAGATACATCTCTAACCCACTAGGAACT 2403
Db 1594 Pro---GlyGlnGlySerGlnTrpValGlyMet----- 1603
Qy 2404 AACACCTGCAGCAGTGTGGTCAGAACATCCCTTATCATGTTCTGAGGCTACACGAGTG 2463
Db 1604 -----AlaLeuGluLeuValGluSerSerProValPheAlaArgArgLeuGlyAspCys 1621
Qy 2464 GCTGACTCGGTG----- 2475
Db 1622 AlaAspAlaLeuAlaProPheValGluTrpSerLeuPheAspValLeuGlyAspGluVal 1641
Qy 2476 -----GTCAACATTGTGTCCCTGCCAGCTGTTAATGCTA----- 2512
Db 1642 AlaIleGlyArgVal-AspValValGlnProValLeuTrpAlaValMetValSerLeuAl 1661
Qy 2513 -----ATCCTTCAGAGTGTCTCCAGGGAAGAGACTAAAGGGA 2550
Db 1661 aGluLeuTrpArgSerPheGlyValValProSerAlaValValGlyHisSerGlnGlyG 1681
Qy 2551 A---GCCACCCAGCGCTCACTGGGAAGTGGACTTCACTGAGTAAAGCCGCTAAATAC 2607
Db 1681 uileAlaAlaCysValAlaGlyAlaLeuThr----- 1692
Qy 2608 GGAAACAAATACCTATTGGTTTTGTAGACACCTTTTCAGATGGGTAGAGCTTATCTCT 2667
Db 1693 -----LeuGluAspGlyAlaArgValValAl 1701

Qy 2668 ACTAAGAAAGAGAGACT-----TCAACCCGTGGTGGCTAAAAA 2702
Db 1701 aLeuArgSerArgAlaLeuLeuAlaLeuSerGlyArgGlyGlyMetValSerValProVa 1721
Qy 2703 AATACTCGAAGAAATTTTTCAGATTTTGAATACCTAAGTAATAGGGTCAGACAATGG 2762
Db 1721 lSerAlaAspArgLeuArgAspArgValGlyLeuSer-----ValAlaAlaValAsnGl 1739
Qy 2763 TCCAGCTTTT-----GTTGCCAGGTAAGTCAGGAGCTG----- 2796
Db 1739 yProAlaSerThrValValSerGlyAlaValGluValLeuGluAlaValLeuAlaGluPh 1759
Qy 2797 -----GCCAAGATATTGGGATTGTGGAACACTGCATTGTGCATACAGACCCCAAG 2849
Db 1759 eProGluAlaAlaArgIleProValAspTyrAlaSerHis----- 1772
Qy 2850 CTGAGCAGGTAGAGAGGATGAATAGAACCATTAAGAGACCTTACTTAATTTAGCCGC 2909
Db 1773 -SerValGlnValGluGlyIleArgGluGlyLeuAlaGluAlaLeuAlaProValArgPr 1792
Qy 2910 GGAGACTGGC-----GTTAATGATTGGATAGC 2936
Db 1792 oArgThrGlyGlnValProPheTyrSerThrValThrGlyArgLeuMetAspThrIleGl 1812
Qy 2937 TCTCTGCCCTTTGTGCTTTTGGGTTAGGAACACCCCTGGACAGTTTGGCTGACCC 2996
Db 1812 uLeuAspAlaGluTyrTrpTyrArgAsnLeuArgGluThrValGluPheGlnSerThrVa 1832
Qy 2997 CTATGAATTTACTTACGGGGA-----CCCCCCCCATTGGT 3032
Db 1832 lGluHisLeuMetArgGlnGlyHisThrValPheValGluAlaSerProHisProValLe 1852
Qy 3033 AGAAAT-----GCTTCTGTACATAGTGTGCTGCTGCTTTTCCACGC 3077
Db 1852 uThrIleGlyValGlnAspThrAlaAspThrThrAspThrAspIleValValThr----- 1870
Qy 3078 TTTGTTCTCTAGGCTCAAGGCACCTTGTGGTGGTGAGACACGCGTGGAGGCAACTCCG 3137
Db 1871 -----GlySerLeuArgArgAspAspGlyThrValGlnArgPheLeuThrSerLeuAl 1888
Qy 3138 GGAG-----GCCTACTCAGAGAGGAGAGA 3161
Db 1888 aGluLeuHisValArgGlyValArgIleAspTrpGlyProLeuPheAlaGlyValSerPr 1908
Qy 3162 CTTGCAGATCCCA---CATGTTTCCAACTGGGAGATTCTAGTCTAGCTAGACCCACCG 3218
Db 1908 oValGluLeuProThrTyrAlaPheGln----- 1917
Qy 3219 TGCAGGAAACCTCGAGACTCGGTGGAAGGGCCCTTATCTGCTACTTTTGACCACCAAC 3278
Db 1918 -----ArgGluArgPheTrpLeuGly-----Al 1925
Qy 3279 GGCTGTGAAGTCCAGGAATCTCCACTCGATCCATGATCCACCTTAACCGCGGCC 3338
Db 1925 aAspIleAlaGluSerAlaValAspThrTrpArgTyrGlnIleSerTrpLysProLeuPr 1945
Qy 3339 ACCTCCCGAT-----TCGGGG---TGGAAAGCC 3363
Db 1945 oAspMetAspProProAlaLeuSerGlyThrTrpLeuAla 1958

RESULT 14

US-10-995-561-659
; Sequence 659, Application US/10995561
; Publication No. US2005027054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24

RESULT 15

```
US-11-109-156-14
; Sequence 14, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-Ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-14

Alignment Scores:
Pred. No.: 0.0262 Length: 745
Score: 126.00 Matches: 134
Percent Similarity: 33.5% Conservative: 77
Best Local Similarity: 21.3% Mismatches: 219
Query Match: 2.0% Indels: 200
DB: 7 Gaps: 32

US-10-723-552-3_COPY_2307_5741 (1-3435) x US-11-109-156-14 (1-745)
QY 805 GACGAACCCCTACACAGACCTGGCCAACTTCAGGATCCAAACCTCAGGTGACCCCTC 864
DB 138 AsnLysIleIleHisArgAspLeu-----LysProGluAsnIleVal 151
QY 865 CTCAGTACGTGGATGACCTGCTTCTGGCGGAGCCCAAAACAGGACTGCTTAGAAGGC 924
DB 152 LeuGlnAspVal-----Gly 156
QY 925 ACGAAGCACTACTGCTGGAATTGCTGACTAGGCTACAGAGCCTCTGCTAAGAAGGCC 984
DB 157 GlyLys---IleIleHisIleIleAspLeuGlyTyrAlaLysAspValAspGlnGly 175
QY 985 CAGATTTCAGGAGAGAGGTAAACATCTTGGGGTACAGATTGGGGACGGGACGCGATGG 1044
DB 176 SerLeuCysThrSerPheValGlyThrLeuGlnTyrLeuAlaProGlu----- 191
```

```
QY 1045 CTGACGGAGGACACGGAAGAAAACCTGTAGTCCAGATACCGGCCCAACACACGCAACAA 1104
DB 192 LeuPheGluAsnLys-----ProTyrThrAlaThrVal 202
QY 1105 ATGAGAGAGTTTTTGGCGACAGCTGGATT---TGCAGACTGTGGATCCCGGGTTGCG 1161
DB 203 AspTyrTrpSerPheGlyThrMetValPheGluCys-----IleAlaGlyTyrArg 219
QY 1162 ACCTTAGCAGCCCACTCTACCCGCTAACCAAGAAAAGGGGAATTCCTCTGGGCTCCT 1221
DB 220 ProPheLeuHisHisLeuGlnPro-----PheThrTrp----- 230
QY 1222 GAGCACCAAGGCAATTTGATGCTATCAAAAAGCCCTGTGTGAGCGCACCTGCTCTGCC 1281
DB 231 ---HisGluLys-----IleLysLys----- 236
QY 1282 CTCCTGACGTAACTAAACCTTTTACCCCTTTATGTG---GATGAGCGTAGGGAGTAGCC 1338
DB 237 -----LysAspProLysCysIlePheAlaCysGluGluMetSerGlyGluVal 252
QY 1339 CGG-----GGAGTTTAAACCCAAACCTA 1362
DB 253 ArgPheSerSerHisLeuProGlnProAsnSerLeuCysSerLeuIleValGluProMet 272
QY 1363 GGACCATGGAGAGACCTGTCCGCTACCTGTCAAAGAAAGCTCGATCTGTAGCCAGTGGT 1422
DB 273 GluAsnTrpLeuGln-----LeuMetLeuAsnTrpAspProGlnGlnArgGly 288
QY 1423 TGGCCCATATGCTGAAGGCTATCGCAGCTGTGGCCATATCGTCAAGGACGCTGACAAA 1482
DB 289 GlyProVal-----Asp 292
QY 1483 TTGACTTTTGGACAG---AATATAACTGTAATAGCCCCCATCATCTGGAGAACATCGTT 1539
DB 293 LeuThrLeuLysGlnProArgCysPheValLeuMetAspHisIleLeu----- 308
QY 1540 CGGCAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCCACTATCAAAAGCTGTT 1599
DB 309 -----AsnLeuLysIleValHisIleLeuAsnMetThr 319
QY 1600 CTCACAGAGAGGTGACGTTTCGCTCCACACGCGCTCTCAACCTCGCACCTTTCTGCT 1659
DB 320 SerAlaLysIleIleSerPheLeuLeuProAspGluSerLeuHisSerLeuGlnSer 339
QY 1660 GAAGAGACTGATGAACCCAGTCACTCATGATTGCCATCACTATTGATTGAGGAGCTGGG 1719
DB 340 ArgIleGluArgGluThrGlyIleAsnThrGlySerGlnGluLeuLeuSerGluThrGly 359
QY 1720 GTC-----CGCAAGGACCTTACAGACATACCGCTGCTGGA----- 1755
DB 360 IleSerLeuAspProArgLysProAlaSerGlnCysValLeuAspGlyValArgGlyCys 379
QY 1756 ---GAAGTGTAACCTGGTTCACTGACGGAAGCAGCTATGTGTGGAAAGGTAAAGAGATG 1812
DB 380 AspSerTyrMetValTyrLeuPheAspLysSerLysThrValTyrGluGlyPro---Phe 398
QY 1813 GCTGGGGCGCGGTGGTGGAC----- 1833
DB 399 AlaSerArgSerLeuSerAspCysValAsnTyrIleValGlnAspSerLysIleGlnLeu 418
QY 1834 -----GGGACCCGACGATCTGGCC-----AGCAGCTCCGC 1866
DB 419 ProIleIleGlnLeuArgLysValTyrAlaGluAlaValHisTyrValSerGlyLeuLys 438
QY 1867 GAAGGAACCTTCAGCAAAAGAGGTGAGCTCATGGCCCTCAGCAAGCTTTGGCGCTGCC 1926
DB 439 GluAspTyrSer-----ArgLeuPhe 445
QY 1927 GAAGGGAATCCATAAACATTTTATACGACAGCAGGTATGCCTTTCGCACTGCACACGTA 1986
DB 446 GlnGlyGlnArgAlaAlaMetLeuSerLeuLeuArgTyr----- 458
```

```
QY 1987 CATGGGCGCATCTATAAACAAGGGGTGCTTACCTCAGCAGGAGGGAATAAAGAAC 2046
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
459 AenAlaAenLeuThrLysMetLysAenThrLysSerAlaSerGlnGlnLeuLysAla 478
QY 2047 AAAGAGGAAATTTCTAAGCCTATTAGACCCGTCATTTTACCAAAAGGCTAGCTATTATA 2106
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
479 Lys-----LeuGluPhePheHisLysSerIleGlnLeuAspLeuGlu 492
QY 2107 CACTGCTCTGAGCATCAG-----AAAGCTAAAGATCTCATATCCAGAGGAAC 2154
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
493 ArgTyrSerGluGlnMetThrTyrGlyIleSerSerGluLysMetLeuLysAlaTrpLys 512
QY 2155 CAGATGGCTGACCGGTTGCCAAGCAGGAGCCAGGGGTGTT----- 2196
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
513 GluMetGluGluLysAlaIleHisTyrAlaGluValGlyValIleGlyTyrLeuGluAsp 532
QY 2197 -----AACCTT---CTGCCTATATAGAAATGCCAAAGCCCCAGAACCCAGACGA 2244
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
533 GlnIleMetSerLeuHisAlaGluIleMetGluLeuGlnLysSerProTyrGlyArgArg 552
QY 2245 CAGTACACCTAGAACTGGCAAGAGATATAAAAGATAGAC-----CAGTTCTCT 2295
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
553 GlnGlyAspLeuMetGluSerLeuGluGlnArgAlaIleAspLeuTyrLysGlnLeuLys 572
QY 2296 GAGACTCCGGAAGGACCTGCTATACCTCAGATGGGAAGGAAATCCTGCCCCACAAAGAA 2355
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
573 HisArgProSerAspHisSerTyr---SerAspSerThrGluMetVal-----LysIle 589
QY 2356 GGGTTAGAATATGTCCAACAGATACATCGTCTAACCCACCTAGGAACCTAAACACCTGCAG 2415
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
590 IleValHisThrValGlnSerGlnAspArgValLeuLysGluLeuPheGlyHisLeuSer 609
QY 2416 CAGTTGGTCAGAACATCCCTTATCATGTTCTGAGGCTACGAGAGTG----- 2463
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
610 LysLeuLeuGlyCysLysGlnLysIleIleAspLeuLeuProLysValGluValAlaLeu 629
QY 2464 -----GCTGACTCGGTGGTC 2478
Db   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
630 SerAenIleLysGluAlaAspAenThrVal 639
```

Search completed: February 14, 2006, 17:24:48

Job time : 255.061 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 15:55:36 ; Search time 9.89233 Seconds
(without alignments)
3199.272 Million cell updates/sec

Title: US-10-723-552-3_COPY_5620_7533

Perfect score: 3468
Sequence: 1 ATGCATCCACGTTAAACCG.....CTAGCAGGAGAGCTGGCCGC 1914

Scoring table:
BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10723552/runat_14022006_125146_13079/app.query.fasta_1
-DB=Issued Patents AA -QPMT=fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abes02p -USER=US10723552 @CGN 1.1 161 @runat_14022006_125146_13079
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/H COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/PCRUS COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3420	98.6	638	2	US-09-376-781-6
2	2866.5	82.7	660	2	US-09-111-085-2
3	2866.5	82.7	660	2	US-09-376-781-5
4	2422	69.8	657	2	US-09-111-085-4
5	2422	69.8	657	2	US-09-376-781-4
6	2395	69.1	657	2	US-09-171-553B-10
7	2353	67.8	656	2	US-09-376-781-3
8	2353	67.8	656	2	US-09-171-553B-6
9	1489.5	42.9	667	2	US-09-315-127-5
10	1489.5	42.9	667	2	US-09-315-127-6
11	1489.5	42.9	667	2	US-09-070-630-13
12	1477.5	42.6	673	2	US-09-075-272-5

13	1319	38.0	665	2	US-09-309-572-14	Sequence 14, Appl
14	1319	38.0	665	2	US-09-718-096-14	Sequence 14, Appl
15	1319	38.0	1312	2	US-09-554-572-26	Sequence 26, Appl
16	1318.5	38.0	632	2	US-09-315-127-2	Sequence 3, Appl
17	1318.5	38.0	632	2	US-09-315-127-3	Sequence 3, Appl
18	1277	36.8	645	2	US-09-315-127-8	Sequence 8, Appl
19	1277	36.8	645	2	US-09-315-127-9	Sequence 9, Appl
20	1262.5	36.4	654	2	US-09-315-127-11	Sequence 11, Appl
21	1262.5	36.4	654	2	US-09-315-127-12	Sequence 12, Appl
22	692.5	20.0	469	1	US-08-484-126-1	Sequence 1, Appl
23	692.5	20.0	469	2	US-09-374-909-1	Sequence 1, Appl
24	659	19.0	445	1	US-08-447-925-6	Sequence 6, Appl
25	639.5	18.4	453	1	US-08-484-126-3	Sequence 3, Appl
26	639.5	18.4	453	2	US-09-374-909-3	Sequence 3, Appl
27	626.5	18.1	196	1	US-08-484-126-7	Sequence 7, Appl
28	626.5	18.1	196	2	US-09-374-909-7	Sequence 7, Appl
29	448	12.9	567	1	US-08-007-282B-2	Sequence 2, Appl
30	414.5	12.0	538	2	US-09-175-928-4	Sequence 4, Appl
31	413.5	11.9	540	2	US-09-949-016-11511	Sequence 11511, A
32	413.5	11.9	540	2	US-09-719-554-26	Sequence 26, Appl
33	413.5	11.9	685	2	US-09-719-554-35	Sequence 35, Appl
34	369	10.6	493	2	US-08-979-847B-106	Sequence 106, App
35	356.5	10.3	192	2	US-08-486-099-107	Sequence 107, App
36	356.5	10.3	192	2	US-08-360-107A-117	Sequence 117, App
37	356.5	10.3	192	2	US-08-484-223B-107	Sequence 107, App
38	356.5	10.3	192	2	US-08-919-597-107	Sequence 107, App
39	356.5	10.3	192	2	US-08-475-668A-107	Sequence 107, App
40	356.5	10.3	192	2	US-08-485-551A-107	Sequence 107, App
41	356.5	10.3	192	2	US-08-471-913A-107	Sequence 107, App
42	356.5	10.3	192	2	US-08-485-264A-107	Sequence 107, App
43	356.5	10.3	192	2	US-08-474-349A-107	Sequence 107, App
44	356.5	10.3	192	2	US-08-470-896-107	Sequence 107, App
45	356.5	10.3	192	2	US-08-485-546A-107	Sequence 107, App

ALIGNMENTS

RESULT
US-09-376-781-6
; Sequence 6, Application US/09376781
; Patent No. 6261806
; GENERAL INFORMATION:
; APPLICANT: Banerjee, Papia T.
; APPLICANT: Patience, Clive
; APPLICANT: Andersson, Goran K.
; TITLE OF INVENTION: Molecular Sequences of Swine Retrovirus and Methods of
; Patent No. 6261806
; TITLE OF INVENTION: Use
; FILE REFERENCE: 61750-267
; CURRENT APPLICATION NUMBER: US/09/376,781
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/097,015
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PERV-C
; OTHER INFORMATION: polypeptide sequence taken from GenBank Accession
; OTHER INFORMATION: No. 6261806 AF038600 for comparison.
US-09-376-781-6

Alignment Scores:
Pred. No.: 0 Length: 638
Score: 3420.00 Matches: 637
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 98.6% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-376-781-6 (1-638)

```
QY 1 ATGCATCCACGTTTAAACCGGGCCACCTCCCGATTCCGGGTGGGAAAGAGACTG 60
Db 1 MetHisProThrLeuAsnArgArgHisLeuProIleArgGlyGlyProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCCATCGCTGGTTCCTTACTCTGTCTCAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGGTATCGCATAGGACAGCCTGAACTCCCATAAACCTTATCTCTCACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCGGCACAGGTATTAATATCAACACACTCAAGGGAGGCTCTTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGTACTATACCTTTGCTCCCTCAGATCAGTTATTCTTCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuThrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCCTCCATCTGCTCACGGATTTTATGTTTCCCGAGGACCAACCAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheThrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGGAATCCAGAGATTTCTTTTGTAAACAATGGAACCTGTGTAACCTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTCGAATGCGCAACCTCTCAGCAGGATAGGTAAGTTTCTTATGTCTCAAC 480
Db 141 AspGlyThrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerThrValAsn 160
QY 481 ACCTATACCGACTCTGGCAATTAATTAACCTGACCTCGATTAAGCTTGAAGCCCAAG 540
Db 161 ThrThrThrSerSerGlyGlnPheAsnThrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCTTCAGACTGATTAACCTAAATAATAGTTTCACTGAGAAAGGAAACAAGAA 600
Db 181 CysSerProSerAspLeuAspThrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTTAATAGGTAAATGGTATGCTTGGGGAATGGTATATTAAGGAGCTCGGGT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValThrTrpGlyLysSerGly 220
QY 661 AAACAACAGGCTCCATTCCTAATCTGCTCAAAATAAACAGCTGGAGCTCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGGACCAATACGGTCTTGAGCGGTCAAGACCCCAACCCCAAGGACCGAGCCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTTCCTGGATCAGACCCCTAGTCTTAACAGCACCACTAAATAGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerSerThrThrLysMetGly 280
QY 841 GCAAAACCTTTTAGCTCATCAGGAGCTTTTCAAGCTCTTAACTCCACGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACTCTTCTTGTGGTATGCTTAGCTTCGGGCCCACTTACTATGAAGGAATGGCT 960
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProThrTrpGlyMetAla 320
QY 961 AGAAGGGAATTCATGTGCAAAAGACATAGACCAATGACATGGGATGCCAA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGTTTCTGGAAGGACCTGTCATAGGAAGGATCCCA 1080
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
```

```
QY 1081 TCCACCAACACCTTTGTAAACCACTGAAGCCCTTAAATCAAACTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnThr 380
QY 1141 CTGTACTCTGTTATGACAGGTGGGCATGTAACTACTGTGATTAAACCCCTTGTGTTC 1200
Db 381 LeuValProGlyThrAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTAAACCAAACTAAAGATTTTTCGATTATGTCCTCAAAATGTTTCCCGAGTG 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAGCAATCTTGATGAATATGACTACAGAAATCATCGCAAAAAG 1320
Db 421 TyrThrThrProGluLysAlaIleLeuAspGluThrAspThrArgAsnHisArgGlnLys 440
QY 1321 AGAGAACCAATATCTCTGACACTTGTCTGTGATGTCGACTTGGAGTGGCAGCAGGTGA 1380
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGAACAGCTGCCCTGGTCACGGACCAACAGAGCTAGAAACAGGACTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTTGTAACAGAGATCTCCAGCCCTAGAAATCTGTCAAGTAACCTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTAACTCTTATCTGAAGTAGTCTCAAGAAATAGAGAGGCTTAGATTTATTA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGATTAATGTAGCTTGAAGGAGGAAATGCTGTTTATGTGGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheThrValAsp 540
QY 1621 CATTCAGGGCCATCAGAGACTCCATGACAGCTTAGAGAAAGGTTGGAGAGGCTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGAAAAGGAAACTACTCAAGGGTGGTTTCAGGGATGGTTTCAACAGGCTCTTTGGTGTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
QY 1741 GCTACCTTACTTCTGCTTAAACAGGACCTTAATAGTCTCTCTCTCTCTTACTACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATCTATTATTAAACAGTTAATTCCTTCATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgGlyLeuAlaValGln 620
QY 1861 ATCATGTGCTTACAGCAACAGTACCAAGCCCGCTCTAGCAGGAGGCTGGCGCG 1914
Db 621 IleMetValLeuArgGlnGlnThrGlnSerProSerSerArgGluAlaGlyArg 638
```

RESULT 2

```
US-09-111-085-2
; Sequence 2, Application US/09111085
; Patent No. 610034
; GENERAL INFORMATION:
; APPLICANT: Stoye, Jonathan P
; APPLICANT: Weiss, Robin A
; TITLE OF INVENTION: Detection of retroviral subtypes based upon envelope
; TITLE OF INVENTION: specific sequences
; FILE REFERENCE: 4238/75168
; CURRENT APPLICATION NUMBER: US/09/111,085
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: GB 9710154.7
; EARLIER FILING DATE: 1997-05-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.0
```

; SEQ ID NO 2
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Porcine retrovirus
US-09-111-085-2

Alignment Scores:

Pred. No.: 7,72e-281 Length: 660
Score: 2866.50 Matches: 537
Percent Similarity: 88.6% Conservative: 39
Best Local Similarity: 82.6% Mismatches: 53
Query Match: 82.7% Indels: 21
DB: 2 Gaps: 6

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-111-085-2 (1-660)

QY 1 ATGCATCCACGTTAAACCGCGCCACCTCCCGATTGGGTGGAAAGCCGAAAGACTG 60
DB 1 MethisProThrLeuSerArgArgHisLeuProIleArgGlyGlyProLysArgLeu 20
QY 61 AAAATCCCTTAAAGCTTCGCTCCATCGCGTTCCTTACTCTGTCATAACCTCTCAG 120
DB 21 LysileProLeuSerPheAlaSerileAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATGGTATGCGCATAGGAGACAGCTGAACCTCCCATAAACCTTATCTCTCACCTGG 180
DB 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGGCACAGGTATTAAATCAACACACTCAAGGGGAGGCTCCTTTA 240
DB 61 LeuLeuThrAspSerGlyThrGlyLeuAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGGCTGATCTATACGTTTGCTCAGATCAGTTATTCCTAGTCTG----- 294
DB 81 GlyThrTrpTrpProGluLeuTy-VaLysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 295 ---ACCTCACCCACAGATATCTCCATGCTCAGGATTTATGTTGGCCAGGACCA 351
DB 101 GlnAlaThrProProAspValLeuArgAlaTyGlyPheTyValCysProGlyProPro 120
QY 352 AATAATGGAACCAATTCGGAAATCCACAGATTTCTTTGTAACCAATGGAACGTGTA 411
DB 121 AsnAsnGluGluTyCysGlyAsnProGlnAspPheCysLysGlnTrpSerCysIle 140
QY 412 ACCTCTAATGATGATATTGAAATGGCCAACTCTCAGCAGGATAGGTAAGTTTCT 471
DB 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnAspArgValSerTySer 160
QY 472 TATGTCACACCTATACAGCTCTGGACAATTTAATTAC-----CTGACC 516
DB 161 PheValAsnAsnProThrSerTyAsnGlnPheAsnTyArgLysGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGACTGGAGCCCAAGTCTCTCCTTCAGAC 555
DB 181 TrpGlnGlnArgValGlnLysAspVaLargAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 556 CTAGATTACCTTAAATAAGTTTCACCTGAGAAAGGAAACAAAGAAATATCTCTAAATGG 615
DB 201 LeuAspTyLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrp 220
QY 616 GTAAATGGTATGCTTCGGGAATGGTATATATGAGGCTCGGGTAAACACCAAGGCTCC 675
DB 221 ValAsnGlyIleSerTrpGlyIleValTyTyGlyGlySerGlyArgLysLysGlySer 240
QY 676 ATTCTAATCTTCGCTCAAAATA- -AACACGCTGGAGCTCAATGGCTATAGGACCA 732
DB 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProProValAlaIleGlyPro 260
QY 733 AATACGCTTTGAGGGTCAAGACCCCAACCCCAA-----GGACACGAGCA 780
DB 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGlnArgProSerProAsnPro 280
QY 781 TCCTCT---AACATACTTCTGGATCAGACCCCACTGAGTCTAACACGACTAAATG 837

Db 281 SerAspTyAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY 838 GGGCAAAACCTTTTAGCTCATCCAGGAGCTTTCAAGCTCTTAACCTCCACGACTCCA 897
DB 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrPro 320
QY 898 GAGCTACCTCTCTTGTGTGGCTATGCTTTCGGGCCCACTTACTATGAAGGAATG 957
DB 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyTyGlyGlyMet 340
QY 958 GCTAGAAGAGGGAATTCATGTCACAAACAAATAGAGACCACATGACATGGGATCC 1017
DB 341 AlaArgGlyGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 1018 CAAATAAGCTTACCTTACTGAGGTTTCTGAAAAGGCCTCATAGAGAAAGTCTCC 1077
DB 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380
QY 1078 CCATCCCAACCAACCTTTGTAAACCAACCTTAAATCAAAACCTCTGAGAGTCAA 1137
DB 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 1138 TATCTGTACTGTTATGACAGGTGGTGGCATGTAATCTGATTAACCCCTTGTGTT 1197
DB 401 TyrLeuValProGlyTyAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 1198 TCCACCTTGGTTTTTAACCAACCTTAAAGATTTTGTGCTTATGCTCCAAATGTTCCCGA 1257
DB 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 1258 GTGTATTACTATCCCGAAAAGCAATCCTTCATGAATATGACTACAGAAATCATCGACAA 1317
DB 441 ValTyTyTyTyProGluLysAlaValLeuAspGluTyAspTyArgTyAsnArgPro 460
QY 1318 AAGAGAAACCATATCTCTGACACTTGTGCTGATGCTCGACTTGAGTGGCAGCAGGT 1377
DB 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
QY 1378 GTAGAACACGACAGCTGCCCTGCTCAGCGGACCAAGCAGCTAGAGAAACAGGACTAGT 1437
DB 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuLysGlyLeuSer 500
QY 1438 AACCTACATCGAATTTGTACAGAGATCTCCAAAGCCCTAGAGAAATCTGTCAAGTAACCTG 1497
DB 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 1498 GAGGAATCCCTAACCTCTTATCTGAAGTGTCTCTACAGAAATAGAGAGGTTAGATT 1557
DB 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 1558 TTATTTCTAAAGAGGAGGATTTATGTTAGCTTGAAGGGAATCTGTTTTTATGTG 1617
DB 541 LeuPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGlyCysPheTyTyVal 560
QY 1618 GATCATTTCAGGGGCCATCAGAGACTCCATGAACAAAGCTTAGAGAAAGTTGGAGAAGCGT 1677
DB 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 580
QY 1678 CGAAGGAAAGGAAACTACTCAAGGTTGTTGAGGGATGTTTCAACAGGTCTCTTTGG 1737
DB 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 1738 TTGCTACCTCTTCTGCTTTAAACAGGACCTTAAATAGTCTCTCTCTCTCTCTCTCTCT 1797
DB 601 MetThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThr 620
QY 1798 GTTGGGCCATGTATTATTAAAGTTAAATGCTCTTATTAGAGAACGAATAAGTGCACTG 1857
DB 621 ValGlyProCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 640
QY 1858 CAGATCATGTACTTTAGACCAACAGTACCA 1887

Db 641 GlnIleMetValLeuArgGlnGlnTyrGln 650

RESULT 3

US-09-376-781-5

; Sequence 5, Application US/09376781

; Patent No. 6261806

GENERAL INFORMATION:

; APPLICANT: Banerjee, Papiya T.

; APPLICANT: Patience, Clive

; APPLICANT: Andersson, Goran K.

; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of

; Patent No. 6261806

; TITLE OF INVENTION: Use

; FILE REFERENCE: 61750-267

; CURRENT APPLICATION NUMBER: US/09/376,781

; CURRENT FILING DATE: 1999-08-18

; EARLIER APPLICATION NUMBER: 60/097,015

; EARLIER FILING DATE: 1998-08-18

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 660

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: PERV-A

; OTHER INFORMATION: Polypeptide sequence taken from GenBank Accession

; OTHER INFORMATION: No. 6261806 Y12238 for comparison.

US-09-376-781-5

Alignment Scores:

Pred. No.:	Length:	Score:
7, 728-281	660	2866.50
Percent Similarity:	Matches:	537
88.6%	Conservative:	39
Best Local Similarity:	Mismatches:	53
82.7%	Indels:	21
Query Match:	Gaps:	6
DB:		

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-376-781-5 (1-660)

QY 1 ATGCATCCACGCTTAAACGGCGCCACCTCCCGATTTCGGGGTGGAAAGCCGAAAGACTG 60

Db 1 MethHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyProLysArgLeu 20

QY 61 AAAATCCCTTAAAGCTTCGCTCCATCGCGGTTCCTTACTCTCTCAATTAACCTCTCAG 120

Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40

QY 121 ACTAATGGTATCGCATAGGACAGCCTGAACTCCCATAAACCTTATCTCTCACCTGG 180

Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60

QY 181 TTAATTAAGTCTCCGCGCACAGTATTAAATCAACAACTCAAGGGAGGCTCTTTTA 240

Db 61 LeuLeuThrAspSerGlyThrGlyLeuAsnIleAsnSerThrGlnGlyGluAlaProLeu 80

QY 241 GGAACCTGGTGGCTGATCTATACGTTGCTCCTCAGATCAGTATTTCCTAGTCTG----- 294

Db 81 GlyThrTrpTrpProGluLeuThrValCysLeuArgSerValIleProGlyLeuAsnAsp 100

QY 295 ---ACCTCACCCAGATATCTCCATGCTCAGGATTTTATGTTGTCAGGACCA 351

Db 101 GlnAlaThrProProAspValLeuArgAlaThrGlyPheThrValCysProGlyProPro 120

QY 352 AATAATGGAACAATTCGGGAATCCAGAGATTTCTTTTGTAAACAATGGAACCTGTGTA 411

Db 121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysLysTrpSerCysIle 140

QY 412 ACCTCTAATGATGGATATTTGGAAATGGCCAACTCTCAGCAGGATAGGTTAGTTTCT 471

Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTrpSer 160

QY 472 TATGTCACACCATATACACGAGCTCTCGACATTTTAATTAC-----CTGACC 516

Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTriPlysAsp 180

QY 517 TGG-----ATTAGAACTGGAAGCCCCAAAGTGTCTCTCTTCAGAC 555

Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200

QY 556 CTAGATTACCTAAAATAAGTTTCACCTCAGAAAGGAAACCAAGAAATATCTTAAATGG 615

Db 201 LeuAspTyrLeuLysIleSerPheThrGlnLysGlyGlnGluAsnIleGlnLysTrp 220

QY 616 GTAAATGGTATCTTGGGAAATGATATTATGGAGGCTCGGTAAACCAACAGGCTCC 675

Db 221 ValAsnGlyIleSerTrpGlyLeuValTyrGlyGlySerGlyArgLysLysGlySer 240

QY 676 ATTCTAACTATTCGCTCAAAATA---AACAGCTGGAGCTTCCAATGGCTATAGGACA 732

Db 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProProValAlaIleGlyPro 260

QY 733 AATACGGTCTTGACGGGTCAAAGACCCCAACCCAA-----GGACCAAGGACCA 780

Db 261 AsnLysGlyLeuAlaGluGlnGlyProIleGlnGlnArgProSerProAsnPro 280

QY 781 TCCTCT---AACATAACTTCTGGATCAGACCCCACTGAGTCTAAACAGCACACTAAAATG 837

Db 281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300

QY 838 GGGGCAAACTTTTACGCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCACGACTCA 897

Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320

QY 898 GAGGCTACCTCTTCTGTTGGCTATGCTTACGCTCGGGCCCACTTACTATGAGGAATG 957

Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrGlnGlyMet 340

QY 958 GCTAGAGAGGAGAAATTCATGTGCACAAAGAAACATAGAGACCAATGACATGGGATCC 1017

Db 341 AlaArgGlyGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360

QY 1018 CAAATAAGCTTACCTTACCTGAGGTTCTCGAAAGGCACCTGATAGGAAGGTTCC 1077

Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380

QY 1078 CCATCCCAACCAACCTTTGTAACACACTCAAGCCTTTAATCAACCTCTGAGAGTCAA 1137

Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGlnSerGln 400

QY 1138 TATCTGGTACCTGGTTATGACAGGTGGTGGGCATGTATATCTGGATTAAACCCCTTGT 1197

Db 401 TyrLeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 420

QY 1198 TCCACCTTGGTTTAAACCAAACTAAAGATTTTTCATTTGCTGATTTGTTCCCGA 1257

Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440

QY 1258 GTGTATTACTATCCGAAAGCAATCCTTCATGAATATGACTACAGAAATCATCGACAA 1317

Db 441 ValTyrTyrProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460

QY 1318 AAGAGAGAACCATATCTCTGACACTTCTGCTGATGCTCGGACTTTGGAGTGGCAGAGT 1377

Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480

QY 1378 GTAGAAACAGAAACGCTGCTGTCAGGGACACAGCAGCTAGAAACAGGACTTAGT 1437

Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500

QY 1438 AACCTACATCGAATTTGTAACAGAAAGCTCTCAAGCCCTAGAAAATCTGTAGTAACTG 1497

Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520

QY 1498 GAGGAATCCCTAAACCTCTTATCTGAAGTAGTCTCTACAGAATAGAAAGGGTGTAGATT 1557

```
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
Qy 1558 TTATTTCTAAAGAGGAGGATTATGTAGCTTGAAGAGGAGGAAATGCTGTTTTATGTG 1617
Db 541 LeuPheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrVal 560
Qy 1618 GATCATTCAGGGCCCATCAGAGACTCCATGAAACAGCTTAGAGAAAGTTTCGAGAGCGT 1677
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArg 580
Qy 1678 CGAAGGGAAGAAAGAACTACTCAAGGGTGGTTTCAGGGATGGTTCAACAGGTCTCTTGG 1737
Db 581 ArgArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyr 600
Qy 1738 TTGGCTACCCCTACTTCTGCTTTTAAACAGGACCTTAATAGTCTCTCTCTTACTCACA 1797
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 620
Qy 1798 GTTGGGCCATGTATTTAAACAAGTTAATGCTCTTATTAGTCTCTCTCTCTCTCTCT 1857
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
Qy 1858 CAGATCATGTACTTAGACACAGTACCACA 1887
Db 641 GlnIleMetValLeuArgGlnGlnTyrGln 650

RESULT 4
US-09-111-085-4
; Sequence 4, Application US/09111085
; Patent No. 6100034
; GENERAL INFORMATION:
; APPLICANT: Stoye, Jonathan P
; APPLICANT: Weiss, Robin A
; TITLE OF INVENTION: Detection of retroviral subtypes based upon envelope
; TITLE OF INVENTION: specific sequences
; FILE REFERENCE: 4238/75168
; CURRENT APPLICATION NUMBER: US/09/111.085
; CURRENT FILING DATE: 1998-07-07
; EARLIER FILING DATE: 1997-05-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Porcine retrovirus
US-09-111-085-4

Alignment Scores:
Pred. No.: 7.14e-236 Length: 657
Score: 2422.00 Matches: 457
Percent Similarity: 80.2% Conservative: 68
Best Local Similarity: 69.8% Mismatches: 96
Query Match: 69.8% Indels: 34
DB: 2 Gaps: 9

US-10-723-552-3_copy_5620_7533 (1-1914) x US-09-111-085-4 (1-657)
Qy 1 ATGCATCCAGCTTAACCCGCGCCACTCCCGATTCGGGGTGAAGAGCGAAAGACTG 60
Db 1 MethisProthrLeuSerTyrArgHisLeuProthrArgGlyGlyGluProLysArgLeu 20
Qy 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG 120
Db 21 ArgIleProLeuSerPheAlaSerIleAlaTyrPheLeuThrLeuThrIleThrProGln 40
Qy 121 ACTAATGGTATGCGCATPAGGAGACAGCTCGAATCCCATAAACCTTTACTCTCACTGG 180
Db 41 AlaSerLysArgLeuIleAspSerSerAsnProHisArgProLeuSerLeuThrTyr 60
Qy 181 TTAATTAAGTACCTCCGCGACAGGTATTAATATCAACAACTCAAGGGAGGCTCTTTA 240
Db 61 LeuIleIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
```

```
Qy 241 GGAACTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTTATCTAGTCTGACCTCA 300
Db 81 GlyThrTyrTrpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
Qy 301 ---CCCCAGATATCTCTCCATGCTCAGGATTTATGTTGCCAGGACACCAATAAT 357
Db 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysCysProGlyThrGluLysGlu 120
Qy 358 GGAAAAATTCGCGAAATCCAGAGATTCTTTTGTAAACAAATGGAACCTGTGAACCTCT 417
Db 121 ---LysTyrCysGlyGlySerGlyGluSerPheCysArgArgTyrPheCysValThrSer 139
Qy 418 AATGATGGATATTCGAAATGCCAACCTCTCAGCAGGATAGGTAAGTTTCTTATGTC 477
Db 140 AsnAspGlyAspTyrLysTrpProIleSerLeuGlnAspArgValLysPheSerPheVal 159
Qy 478 AACACCTATACCACTCTCGGACAAATTAATTAATCACTGACCTGGATTAGAACTGGAAGCCCC 537
Db 160 Asn-----SerGlyProGlyLysTyrLysValMetLysLeuTyrLysAspLys 175
Qy 538 AAGTGCTCTCTTCAGACCTAGATTACCTAAATAAGTTTCTCAGAGAAAGGAAACAA 597
Db 176 SerCysSerProSerAspLeuAspTyrLysLysIleSerPheThrGluLysGlyLysGln 195
Qy 598 GAAAAATATCTAAATGGTAAATGGTATGTCTTGGGGAATGGGTATAT-----TATGCA 651
Db 196 GluAsnIleGlnLysTyrPheAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGly 215
Qy 652 GGCTCGGGTAAACAAACCAAGCTCCATTTCTAATTCGCTCAAAATAAACACAG---CTG 708
Db 216 GlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThr 232
Qy 709 GAGCTCCATGCTATAGGACCAATACCGTCTTGACGGGTCAAGACCCCAACCCAA 768
Db 233 GluProProValAlaValGlyProAspLysValLeuAlaGluGlnGlyProProAlaLeu 252
Qy 769 GGACCA-----GGA 777
Db 253 GluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGln 272
Qy 778 CCATCTCTTAACATAACTTCTGGATCAGACCCCACTGAGTCT-----AACAGC----- 825
Db 273 ProProSerAsnGlyThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 292
Qy 826 ---ACGACTAAATGGGGCAAACTTTTATGCTCATCCAGGAGCTTTTCAAGCTCTT 882
Db 293 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 312
Qy 883 AACTCCACGACTCCAGAGGCTACTCTTCTTGTGGTATGCTTAGCTTCGGGCGCCACCT 942
Db 313 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 332
Qy 943 TACTATGAAGAAATGGCTAGAAGAGGAAATCAATGTGACAAAGAACATAGAGACCAA 1002
Db 333 TyrTyrGluGlyMetAlaLysGluGlyLysPheAsnValThrLysGluHisArgAsnGln 352
Qy 1003 TGCACATGGGGATCCCAAAATAAGCTTACCTTACTGAGGTTTCTGAAAGAGGACCTGC 1062
Db 353 CysThrTrpGlySerArgAsnLysLeuThrGluValSerGlyLysGlyThrCys 372
Qy 1063 ATAGGAAGGTTCCCCCATCCCAACACCTTTGTAAACCACTGAAGCTTAAATCAA 1122
Db 373 IleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 392
Qy 1123 ACCTCTGAGAGTCAATATCTGGTACCTGGTTATGACAGAGTGGTGGGCATGTAAATCA 1182
Db 393 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGly 412
Qy 1183 TTAACCCCTTGTGTTCACCTTGGTTTAAACCAAACTAAAGATTTTTCATTTATGCTG 1242
Db 413 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMetVal 432
```

	Qy	1243	CAAAATGTTCCCGAGGTATTACTATCCGAAAAAAGCAATCCTTGATGAATATGACATC	1302
	Dd	433	GlnIleValProArgValTyrTyHisProGluGluValValLeuAepGluTyrAspTyr	452
	Qy	1303	AGAAATCATCGACAARAAGAGAGAACCCATATCTCTGACATCTTCGTGTGATGCTCGGACTT	1362
	Dd	453	ArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu	472
	Qy	1363	GGAGTGCAGCAGGTGTAGGAACAAGCAAGCTGCCCTGGTCTCGGGACCACACAGCAGCTA	1422
	Dd	473	GlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu	492
	Qy	1423	GAAPACAGACTTAGTAACCTACATCGAAATGTATAACAGAAGATCTCCAGGCCCTAGAAAA	1482
	Dd	493	GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuGluGlu	512
	Qy	1483	TCTGTCTAGTAACCTGGAGGAATCCTAACTCCTTATCTCAAGTAGTCCTACAGAATAGA	1542
	Dd	513	SerValSerAsnLeuGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg	532
	Qy	1543	AGAGGTTAGATTATTATTCTTAAAGAAGAGGAGTATATGTGTAGCCTTTGAAGGAGAA	1602
	Dd	533	ArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluGlu	552
	Qy	1603	TGCTGTTTTTTATGTGGATCATTCAGGGCCATCAGACACTCCATGAACRAGCTTAGAGAA	1662
	Dd	553	CysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGlu	572
	Qy	1663	AGGTTGGAGAAGCGCTCGAAGGGAAAAGAACTACTCAAGGGTGGTTTGAGGAGTGGTTC	1722
	Dd	573	ArgLeuGluArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe	592
	Qy	1723	AACAGTCTCTTTGGTTGGCTACCCCTACTTTCTGCTTTTAACAGACCCTTAATAGTCCTC	1782
	Dd	593	AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeu	612
	Qy	1783	CTCCCTGTTACTCACAGTTGGGCCATGTATTATTAAACAAGTTAATGTCTTCATTAGAGAA	1842
	Dd	613	LeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlu	632
	Qy	1843	CGAATAAGTGCAGTCCAGATCATCGTTACTTAGACAACAGTACCAA	1887
	Dd	633	ArgValSerAlaValGlnIleMetValLeuArgGlnGlnTrvGln	647

RESULT 5

```

US-09-376-781-4
; Sequence 4, Application US/09376781
; Patent No. 6261806
; GENERAL INFORMATION:
; APPLICANT: Banerjee, Papia T.
; APPLICANT: Pattence, Clive
; APPLICANT: Andersson, Goran K.
; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
; Patent No. 6261806
; TITLE OF INVENTION: Use
; FILE REFERENCE: 61750-267
; CURRENT APPLICATION NUMBER: US/09/376,781
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/097,015
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PERV-B
; OTHER INFORMATION: polypeptide sequence taken for comparison from
; OTHER INFORMATION: GenBank Accession No. 6261806 Y12239.
US-09-376-781-4
Alignment Scores:

```



```
QY 883 AACTCCAGGACTCCAGAGGCTACCTCTCTTCTGTTGGCTATGCTTAGCTTCGGGCGCCACCT 942
Db 313 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 332
QY 943 TACTATGAAGGAATGGCTAGAGAGGGAATTCATGTGACAAAGACACAGACACAA 1002
Db 333 TyrTrpGluGlyMetAlaGlyGluGlyPheAsnValThrLysGluHisArgAsnGln 352
QY 1003 TGCACATGGGGATCCCAATAAGCTTACCCCTTACTGAGGTTCCTGGAAGGACGACCTGC 1062
Db 353 CysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCys 372
QY 1063 ATAGAAAGGTTCCCCCATCCACACACCTTTGTATACACACACTGAAGCCCTTAAATCAA 1122
Db 373 IleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrValValThrGluGln 392
QY 1123 ACCTCTGAGAGTCAATATCTGCTACCTGCTTATGACAGGTGGTGGCATGTAATACCTGA 1182
Db 393 AlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGly 412
QY 1183 TTAACCCCTTGTTTCCACCTTGCTTTTAAACCAAACTAAAGATTTTTCATTATGCTC 1242
Db 413 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMetVal 432
QY 1243 CAATTTGTTCCCGAGTGATTACTATCCCGAAAAAGCAATCCTTGATGATATGACTAC 1302
Db 433 GlnIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAspTyr 452
QY 1303 AGAAATCATCGACAAAGAGAGACCCATATCTCTGACACTGCTGTGATGCTCGGACTT 1362
Db 453 ArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu 472
QY 1363 GGAGTGGCAGCAGGTGTAGGAACAGGAACAGCTCCCTGCTCAGCGGACACACAGCTA 1422
Db 473 GlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuThrGlyProGlnLeu 492
QY 1423 GAAACAGGACTTAGTAACCTACATCGAATGTGTAAACAGAGATCTCCAGCCCTAGAAAAA 1482
Db 493 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuGluGlu 512
QY 1483 TCTGTAGTAACCTGGAGGATCCCTAACCTCTTATCTGAGTCTGATGCTCCTACAGATAGA 1542
Db 513 SerValSerAsnLeuGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg 532
QY 1543 AGAGGTTAGATTATTATTCTAAAGAGGAGGATTATGTGTAGCTTGAAGGAGGAA 1602
Db 533 ArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyCysAlaAlaLeuLysGluGlu 552
QY 1603 TGCTGTTTTATGTGGATCATTCAGGGGCCATCAGACACTCCATGAACAAAGCTTATAGAA 1662
Db 553 CysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGlu 572
QY 1663 AGTTGGAGAGCGTCCAGGGAAGAAAGAACTACTCAAGGTGGTTGAGGATGCTTC 1722
Db 573 ArgLeuGluArgArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 592
QY 1723 AACAGGTCTCTTGGTGGCTACCTACTTCTTCTTAAACAGGACCTTAAATAGTCTC 1782
Db 593 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeu 612
QY 1783 CTCTGTGTACTCAGTGTGGCCATGTATTATTAACAAAGTAAATGCTTCATTATAGAA 1842
Db 613 LeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlu 632
QY 1843 CGAATAGTCCAGTCCAGATCATGGTACTTAGACACAGTACCA 1887
Db 633 ArgValSerAlaValGlnIleMetValLeuArgGlnTyrGln 647
```

RESULT 6

US-09-171-553B-10

; Sequence 10, Application US/09171553B

; Patent No. 6756227

; GENERAL INFORMATION:

```
; APPLICANT: GALBRAITH, DANIEL N.
; APPLICANT: HAWORTH, CHRISTINE
; APPLICANT: LEES, GILLIAN M.
; APPLICANT: SMITH, KENNETH T.
; TITLE OF INVENTION: PORCINE RETROVIRUS
; FILE REFERENCE: CFV-5.01
; CURRENT APPLICATION NUMBER: US/09/171.553B
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: PCT/GB97/01087
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: GB 9702668.6
; PRIOR FILING DATE: 1997-02-10
; PRIOR APPLICATION NUMBER: GB 9608164.1
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 657
; TYPE: PRN
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Amino acid sequence
; OTHER INFORMATION: of ENV region "Raji"
US-09-171-553B-10
```

Alignment Scores:

Pred. No.:	3.85e-233	Length:	657
Score:	2395.00	Matches:	451
Percent Similarity:	80.3%	Conservative:	75
Best Local Similarity:	68.9%	Mismatches:	95
Query Match:	69.1%	Indels:	34
DB:	2	Gaps:	10

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-171-553B-10 (1-657)

```
QY 1 ATGCATCCACGTTAAACCGCGCCACCTCCCGATTCGGGTGAAAGCGAAAGACTG 60
Db 1 MetHisProThrLeuSerTrpArgHisLeuProThrArgGlyGlyGluProLysArgLeu 20
QY 61 AAATCCCTTAAGCTTCGGCTCCATCGCTGGTCTCTTACTCTGTCAATAACCTCTCAG 120
Db 21 ArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleThrProGln 40
QY 121 ACTAATGGTATGCGCATAGGAGACAGCTCAACTCCCATAAACCTTATCTCTCACCTGG 180
Db 41 AlaSerSerLysArgLeuIleAspSerSerAsnProHisArgProLeuSerProThrTrp 60
QY 181 TTAATTACTGACTCCGCGCACAGGTATTAAATATCAACAACACTCAAGGGAGGCTCTTTA 240
Db 61 LeuIleIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
QY 241 GGAACCTGGTGGCTGATCTATAGTTTGCCTCAGATCAGTATTCTCTAGCTGACCTCA 300
Db 81 GlyThrTrpTrpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
QY 301 ---CCCCACATATCTCCATGCTCAGGATTTTATGTTTCCAGGACACCAATAAT 357
Db 101 ThrProAsnLeuValArgSerTyrGlyPheTyrCysProGlyThrGluLysGlu 120
QY 358 GGAACAATTCGCGAAATCCAGAGATTTCTTTTGTAAACAATGGAACCTGTGAACCTCT 417
Db 121 ---LysTyrCysGlyGlySerGlyGluSerPheCysArgArgTrpSerCysValThrSer 139
QY 418 AATGATGATATTGGAATGGCCAACTCTCAGCAGATAGGGTAAAGTTTCTTATGTC 477
Db 140 AsnAspGlyAspTrpLysTrpProIleSerLeuGlnAspArgValLysPheSerPheVal 159
QY 478 AACACCTATACAGCTCTGGACAATTTAATTACTGACCTGGATTAGAACTGGAAGCCCC 537
Db 160 Asn-----SerGlyProGlyLysTyrLysMetCysLeuTyrLysAspLysSer--- 176
QY 538 AAGTGTCTCTCTTTCAGACCTTAGATTACCTAAAAAATAGTTTTCACCTGAGAAGGAAACAA 597
Db 538 AAGTGTCTCTCTTTCAGACCTTAGATTACCTAAAAAATAGTTTTCACCTGAGAAGGAAACAA 597
```

```
Db 177 --CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGln 195
Qy 598 GAAATATCCTAAATGGTAAATGGTATGCTCTGGGGAATGGTATAT-----TATGGA 651
Db 196 GluAsnIleGlnLysTrpIleAsnGlyMetSerTrpGlyIleValPheTyrLysTyrGly 215
Qy 652 GGCTCGGTAACAACACAGCGCTCCATTCTAACTATTTCGCCTCAAAATAAACACAG---CTG 708
Db 216 GlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThr 232
Qy 709 GAGCTCCCAATGGCTATAGGACCAATACGGTCTTTGACGGGTCAAGACCCCAACCCAA 768
Db 233 GluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAlaLeu 252
Qy 769 GGACCA-----GGA 777
Db 253 GluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGln 272
Qy 778 CCATCCTCTAAACATAACTCTGGATCGACACCCCTAGTCT-----ACAGC----- 825
Db 273 ProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 292
Qy 826 ---ACGACTAAATGGGGCAAACTTTTAGCCCTCATCCAGGGAGCTTTTCAAGCTCTT 882
Db 293 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 312
Qy 883 AACTCCAGACTCCAGAGGCTACCTCTCTTCTTGGCTATGCTTAGCTTCGGGGCCACCT 942
Db 313 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 332
Qy 943 TACTATGAAGAAATGGCTAGAGGGGAAATTCATGTGACAAAAGACATAGAGACCAA 1002
Db 333 TyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsnGln 352
Qy 1003 TGCACATGGGATCCCAAAATAGCTTACCCTTACTGAGGTTTCTCGAAAAGGACCTGC 1062
Db 353 CysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCys 372
Qy 1063 ATAGAAAGGTTCCCGATCCCAACACACCTTTGTAAACCACTTGAACGACCTTTAATCAA 1122
Db 373 IleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrValValTyrGluGln 392
Qy 1123 ACCTCTGAGATCAATATCTGTACTCTGTTATGACAGGTGGTGGCATGTAATACTGGA 1182
Db 393 AlaserGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAlaCysAsnThrGly 412
Qy 1183 TTAACCCCTTGTTTCCACCTTGTTGTTTTTAAACCAACTAAAGATTTTTCATTATGTC 1242
Db 413 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspLeuCysValMetVal 432
Qy 1243 CAATTTGTTCCCGAGTGTTATCTATCCCGAAAAGCAATCTTGATGAATATGACTAC 1302
Db 433 GlnIleValProArgValTyrTyrHisProGluGluValValLeuAspGluTyrAspTyr 452
Qy 1303 AGAATCATCACAAAAGAGAGAACCCATATCTCTGACACTTGTCTGTCATGCTCGACTT 1362
Db 453 ArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu 472
Qy 1363 GGAGTGGCAGCAGGTGTAGGAACAGGAACAGCTGCGCTGGTCCAGGGACCAACAGACTA 1422
Db 473 GlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeu 492
Qy 1423 GAAACAGACTTAGTAACCTACATCGAATTTGAACAGAAGATCTCCAAGCCCTAGAAAAA 1482
Db 493 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGlu 512
Qy 1483 TCTGTCACTAACCTGAGGAATCCCTTAACCTCTTATCTGAAGTAGTCTCTACAGATAGA 1542
Db 513 SerValSerAsnLeuGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg 532
Qy 1543 AGAGGTTTAGATTTATTTCTAAAGAAAGGAGGATTATGTGTAGCCCTTGAAGAGGAA 1602
Db 533 ArgGlyLeuAspLeuPheLeuArgGluGlyGlyLeuCysAlaAlaLeuLysGluGlu 552
```

```
Qy 1603 TGCTGTTTTTATGTGGATCATTTCAAGGGCCCATCAGAGACTCCATGAACAAGCTTTAGAGAA 1662
Db 553 CysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgLys 572
Qy 1663 AGGTTGGAGAAGCGTCCGAAGGAAAAGAACTACTCAAGGGTGGTTTGAAGGGATGGTTC 1722
Db 573 LysLeuGluArgArgArgGluArgGluAlaAspGlnGlyTyrPheGluGlyTyrPhe 592
Qy 1723 AACAGGCTCTTTTGGTTGGCTACCCCTACTTCTCTCTTTAACAGGACCTTAAATAGTCTC 1782
Db 593 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeu 612
Qy 1783 CTCCTGTTACTTCACAGTTGGGCCCATGTATTATTAAACAAGTTAATTGCCTTCATTAGAGAA 1842
Db 613 LeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAlaPheValArgGlu 632
Qy 1843 CGAATAAGTCAGTCCAGATCATGTACTTAGACAACAGTACCAA 1887
Db 633 ArgValSerAlaValGlnIleMetValLeuArgGlnTyrGln 647

RESULT 7
US-09-376-781-3
; Sequence 3, Application US/09376781
; Patent No. 6261806
; GENERAL INFORMATION:
; APPLICANT: Banerjee, Papia T.
; APPLICANT: Patience, Clive
; APPLICANT: Andersson, Goran K.
; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
; Patent No. 6261806
; TITLE OF INVENTION: Use
; FILE REFERENCE: 61750-267
; CURRENT APPLICATION NUMBER: US/09/376,781
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/097,015
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Amino acid
; OTHER INFORMATION: sequence of a PERV env polypeptide previously
; OTHER INFORMATION: described
US-09-376-781-3

Alignment Scores:
Pred. No.: 6,81e-229 Length: 656
Score: 2353.00 Matches: 450
Percent Similarity: 79.8% Conservative: 74
Best Local Similarity: 68.5% Mismatches: 96
Query Match: 67.8% Indels: 37
DB: 2 Gaps: 10

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-376-781-3 (1-656)
Qy 1 ATGCATCCACGCTTAAACCGCGCCACCTCCCGATTCCGGGTGGAAACCCGAAAGACTG 60
Db 1 MethisProThrLeuSerArgHisLeuProThrArgGlyGlyGluProLysArgLeu 20
Qy 61 AAAATCCCTTAAAGCTTCGCCTCCATCGCGTGGTTCCTTACTCTGTCAATACCTCTCAG 120
Db 21 ArgIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleThrProGln 40
Qy 121 ACTAATGATGTCGATAGGAGACAGCGCTGAACCTCCATAAACCTTATCTCTACCTGG 180
Db 41 AlaserSerLysArgLeuIleAspSerAsnProHisArgProLeuSerLeuThrTrp 60
Qy 181 TTAATTAATCTCGGCACAGAGTATTAAATATCAACAACACTCAAGGGAGGCTCTTTA 240
Db 181 TTAATTAATCTCGGCACAGAGTATTAAATATCAACAACACTCAAGGGAGGCTCTTTA 240
```

1243	Qy	CAAA	TTGTTCCCGAGG	TTA	TACTAT	CCGAAAAAGCAAT	CCCTTGATGAATAT	GACTAC	1302													
432	Db	GlnIle	ValPro	ArgVal	TyTrp	HisPro	GluGlu	ValValLeu	AspGlu	TyrAsp	Tyr	451										
1303	Qy	AGAA	ATCATCG	ACAAAG	ACAG	AGAC	CAAC	CAAT	CTCT	GCAC	ACTT	CGTGTG	ATGATCT	CGGACTT	1362							
452	Db	Arg	Tyr	Asn	Arg	Pro	Arg	Pro	Arg	Val	Ser	Leu	Thr	Leu	Ala	Val	Met	Leu	Gly	Leu	471	
1363	Qy	GGAG	TGGC	CAGC	AGGT	GTG	AGAA	CAGG	AAAC	GAGT	CGCC	TGCT	GCAT	CGGAC	CAC	AGC	AGCTA	1422				
472	Db	Gly	Thr	Ala	Val	Gly	Val	Gly	Thr	Gly	Thr	Ala	Ala	Leu	Ile	Thr	Gly	Pro	Gln	Gln	Leu	491
1423	Qy	GAAC	AGGACT	TAGT	AACT	ACAT	CGAT	TGTA	TGTA	AC	AGAG	AGAT	CTCC	AGG	CCCT	TAG	AAAA	1482				
492	Db	Glu	Leu	Gly	Leu	Gly	Leu	His	Ala	Ala	Met	Thr	Glu	Asp	Leu	Arg	Ala	Leu	Leu	Gly	Asp	511
1483	Qy	TCGT	CTCAGT	AACT	CGG	AGAA	TCC	TAA	CTCT	TAT	CTCA	AGT	TAGT	CTCT	PAC	AGAT	TAGA	1542				
512	Db	Ser	Val	Ser	Asn	Leu	Glu	Glu	Ser	Leu	Thr	Ser	Leu	Ser	Glu	Val	Val	Leu	Gln	Asn	Arg	531
1543	Qy	AGAG	GGTTAG	ATTAT	TAT	TAT	TCT	TAA	AGAG	AGG	AGT	TAT	TGT	TAG	CTCT	TGA	AGG	AGAA	1602			
532	Db	Arg	Gly	Leu	Asp	Leu	Leu	Phe	Leu	Arg	Glu	Gly	Gly	Leu	Cys	Ala	Leu	Leu	Gly	Asp	551	
1603	Qy	TGCT	GTCTTT	TAT	CTG	ATCAT	TTC	CAGG	GGCC	ATC	AGAC	AGT	CCAT	CAAT	GAAC	CAAG	CTTAG	AGAA	1662			
552	Db	Cys	Cys	Phe	Tyr	Val	Asp	His	Ser	Gly	Ala	Ile	Arg	Asp	Ser	Met	Asn	Leu	Arg	Leu	571	
1663	Qy	AGGT	GTGAG	AAGCGT	CGA	AGG	GGAAAA	GAG	AAAT	CTAC	TCA	AGG	GTGGT	TTG	AGG	CACT	TGCTC	1722				
					</																	

RESULTS 8
US-09-171-5538-6

US-09-171-333B-8
: Sequence 6, Application US/09171553B

: Patent No. 6756227

: GENERAL INFORMATION:

APPLICANT: GALBRAITH, DANIEL N.

APPLICANT: HAWORTH, CHRISTINE

APPLICANT: LEES, GILLIAN M.

APPLICANT: SMITH, KENNETH T.

1. TITLE OF INVENTION: PORC

; FILE REFERENCE: CFV-5.01

; CURRENT APPLICATION NUMBER: US/09/171,553B

; CURRENT FILING DATE: 1999-02-08

; PRIOR APPLICATION NUMBER: PCT/GB97/01087

; PRIOR FILING DATE: 1997-04-18

; PRIOR APPLICATION NUMBER: GB 9702668.6

;
;
; PRIOR FILING DATE: 1997-02-10

;; PRIOR APPLICATION NUMBER: GB 9608164.1

PRIOR FILING DATE: 1996-04-19

; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Paton In Ver 3.1; SOFTWARE: PA
; CEO ID NO €

```

; SEQ ID NO 6
: LENGTH: 6

```

```

; LENGTH: 656
; TYPE: PPT

```

ORGANISM: Porcine retrovirus

US-09-171-553B-6

0-DECC-T/T-EN-EN

Alignment Scores:

Alignment Scores:

Pred. No.: 6,81e-229 Length: 656
Score: 2353.00 Matches: 450
Percent Similarity: 79.8% Conservative: 74
Best Local Similarity: 68.5% Mismatches: 96
Query Match: 67.8% Indels: 37
DB: 2 Gaps: 10

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-171-553B-6 (1-656)

QY 1 ATGCATCCAGCTTAAACCGGCGCACCTCCGATTCGGGTGGAAAGCCGAAAGACTG 60
DB 1 MethisProThrLeuSerArgHisLeuProThrArgGlyGlyGluProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCCTCCATCGCGGTGCTTACTCTCTCAATAACCTCTCAG 120
DB 21 ArgileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuThrIleThrProGln 40
QY 121 ACTAATGGTATCGCATAGGAGACAGCTGAACCTCCCATAAACCCCTTATCTCTCACCTGG 180
DB 41 AlaSerSerLysArgLeuIleAspSerSerAsnProHisArgProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGATCCGGCACAGGTATTAATCAACAACACTCAAGGAGGAGCTCTTTA 240
DB 61 LeuIleIleAspProAspThrGlyValThrValAsnSerThrArgGlyValAlaProArg 80
QY 241 GGAACCTGGTGGCTGATACGTTTGCCTCAGATCAGTATTCCTAGTCTGACCTCA 300
DB 81 GlyThrTrpTrpProGluLeuHisPheCysLeuArgLeuIleAsnProAlaValLysSer 100
QY 301 ---CCCCAGATATCCTCATGCTCAGCATTTTATGTTTCCAGGACACCAATAAT 357
DB 101 ThrProProAsnLeuValArgSerTrpGlyPheTrpCysCysProGlyThrGluLysGlu 120
QY 358 GGAAACATATCGGAAATCCAGAGATTCTTTTCTAAACAATGAATGTTTCTTATGTC 417
DB 121 ---LysTrpCysGlySerGlyGluSerPheCysArgArgTrpSerCysValThrSer 139
QY 418 AATGATGATATGGAATGGCCACCTCTCAGCAGGATAGGTAAAGTTTCTTATGTC 477
DB 140 AsnAspGlyAspTrpLysTrpProIleSerLeuGlnAspArgValLysPheSerPheVal 159
QY 478 AACACCTATACGACTCTGGACAATTTAATTACCTGACCTGGATGAGAACGAGCCCC 537
DB 160 Asn-----SerGlyProGlyLysTrpLysMetLysLeuTrpLysAspLysSer--- 176
QY 538 AAGTCTCTCTCAGACTAGATTACTTAAATAAGTTTCACTGAGAAAGGAAACAA 597
DB 177 ---CysSerProSerAspLeuAspTrpLeuLysIleSerPheThr--GluArgLysThrG 195
QY 598 GAAATATCCTAAATGGTAAATGGTATGCTTGGGGAATGGTATAT-----TATGGA 651
DB 195 LysLysTrpSerLysValAspLysTrpTrpGluLeuGlyAsnSer--PheLeuLeuTrpGly 214
QY 652 GGCTCGGTAAACAACACAGGCTCCATTCTAATCTTCGCTCAAAATAAACACGAG---CTG 708
DB 215 GlyGlyAla-----GlySerThrLeuThrIleArgLeuArgIleGluThrGlyThr 231
QY 709 GAGCTCCCAATGGCTATAGGACCAATAACGCTTTGACGGGTCAAAGACCCCAACCAA 768
DB 232 GluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGlyProProAlaLeu 251
QY 769 GGACCA-----GGA 777
DB 252 GluProProHisAsnLeuProValProGlnLeuThrSerLeuArgProAspIleThrGln 271
QY 778 CCATCTCTAAACATAACTTCTGGATCAGACCCCACTGAGTCT-----AACAGC----- 825
DB 272 ProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArgAsnSerProGly 291
QY 826 ---ACGACTAAATCGGCGCAAAACCTTTTACCTCATCCAGGAGCTTTTCAAGCTCTT 882
DB 292 ValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAlaPheGlnAlaIle 311

QY 883 AACTCCAGACTCCAGAGGCTACCTCTTCTTGTGGCTATCTAGCTTCGGGCCCACT 942
DB 312 AsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSerSerGlyProPro 331
QY 943 TACTATGAAGGAATGGCTGAGAAGGAAATTCATATGTGACAAAAGAACATAGAGACCA 1002
DB 332 TyrTrpGluGlyMetAlaLysGluArgLysPheAsnValThrLysGluHisArgAsnGln 351
QY 1003 TGCATGGGATCCCAATAAATACCTTACTGAGGTCTTGGAAAGGACCCCTGC 1062
DB 352 CysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCys 371
QY 1063 ATAGAAAGGTTCCTCCATCCCAACACCTTTGTAAACACACTGAAGCCTTTAATCAA 1122
DB 372 IleGlyLysAlaProProSerHisGlnHisLeuCysTrpSerThrValValTrpGluGln 391
QY 1123 ACCTCTGAGATCAATATCTGGTACCTGTTATGACAGGTGGTGGCATGTAATCTGGA 1182
DB 392 AlaSerGluAsnGlnTrpLeuValProGlyTrpAsnArgTrpTrpAlaCysAsnThrGly 411
QY 1183 TTAACCCCTTGTGTTCCACCTTGGTTTTAACCAACTAAAGATTTTTTCATTTATGTC 1242
DB 412 LeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPheCysValMetVal 431
QY 1243 CAAATTGTTCCCGAGTGTATTACTATCCCGAAAAAGCAATCCTTGATGAATATGACTAC 1302
DB 432 GlnIleValProArgValTrpTrpHisProGluGluValValLeuAspGluTrpAspTrp 451
QY 1303 AGAATCATCCACAAAAGAGAACCCATATCTCTGACACTTGTGCTGTGATGCTCGGACTT 1362
DB 452 ArgTrpAsnArgProLysArgGluProValSerLeuThrLeuAlaValMetLeuGlyLeu 471
QY 1363 GGAGTGGCAGCAGGTGTAGGAACAGGACAGCTCCCTGTCACGGGACACAGCAGCTA 1422
DB 472 GlyThrAlaValGlyValGlyThrGlyThrAlaLeuIleThrGlyProGlnGlnLeu 491
QY 1423 GAAACAGACTTAGTAACTACATCGAATTCGTAACAGAGATCTCCAAAGCCCTAGAAAA 1482
DB 492 GluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArgAlaLeuLysGlu 511
QY 1483 TCTGTAGTAACTCGAGGAATCCCTAACCTCTTATCTGAAAGTAGTCTCTCAGAAATAGA 1542
DB 512 SerValSerAsnLeuGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArg 531
QY 1543 AGAGGTGTAGATTATTTCTAAAGAGAGGATTTATGTTAGCTTCGAAGGAGAA 1602
DB 532 ArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaLeuLysGlu 551
QY 1603 TGCTGTTTTTATGTGGATCATTCAGGGGCCATCAGAGACTCCATGAACCAAGCTTAGAGAA 1662
DB 552 CysCysPheTrpValAspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgLys 571
QY 1663 AGTTTGAAGAGCGTCGAAGGAAAGAAACTACTCAAGGGTGGTTTGGAGGATGGTTC 1722
DB 572 LysLeuGluArgArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPhe 591
QY 1723 AACAGTCTCTTGTGGTACCTTCTTCTTAAACAGGACCCCTTAATAGTCTCTC 1782
DB 592 AsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeu 611
QY 1783 CTCCTGTTACTCACAGTTGGGCTATGTTATTAAAGATTAATTCCTTCAATTAGAGAA 1842
DB 612 LeuLeuLeuLeuThrValGlyProCysLeuLeuAsnArgPheValAlaPheValArgGlu 631
QY 1843 CGAATAAGTCAGTCCAGATCATGTTACTTATAGAACCAACATACCAA 1887
DB 632 ArgValSerAlaValGlnIleMetValLeuArgGlnGlnTrpGln 646

RESULT 9

US-09-315-127-5
; Sequence 5, Application US/09315127
; Patent No. 6448390
; GENERAL INFORMATION:

```

; APPLICANT: The University of Tennessee, c/o Richard Cox
; TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and
; FILE OF INVENTION: Liposome Vectors and Use in Gene and Drug Therapy
; REFERENCE: 44137-5023, U. of Tennessee
; CURRENT APPLICATION NUMBER: US/09/315,127
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-315-127-5

Alignment Scores:
Pred. No.: 1,55e-141 Length: 667
Score: 1489.50 Matches: 314
Percent Similarity: 58.0% Conservative: 90
Best Local Similarity: 45.1% Mismatches: 171
Query Match: 42.9% Indels: 121
DB: 2 Gaps: 20

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-315-127-5 (1-667)

QY 10 ACCTTAACCCGGCCACCTC-----CCGATTGGGGTGGAAGCCGAAAGACTGAAA 63
Db 11 ThrSerAsnLeuHisHisLeuArgHisGlnMetSerProGlySerTrpLysArgLeu--- 29
QY 64 ATCCCTTAAGCTTCGCGCTCCATCGCGTGGTTCCTTACTCTGCTCAATAACCTCTCAGACT 123
Db 30 -----llelleuLeuSerCysValPheGlyGly 39
QY 124 AATGGTATGCGATAGAGACAGAGCTGAATCCCAATAACCTTATCTCTCACCTGGTTA 183
Db 40 GlyGlyThrSerLeu---GlnAsnLysAsnProHisGlnPrometThrLeuThrTrpGln 58
QY 184 ATTACTGACTCCGGACAGGATTAAATATCAACACACTCAAGGGGAGGCTCCTTTAGGA 243
Db 59 ValLeu---SerGlnThrGlyAspValTrpAspThrLysAlaValGlnProProTrp 77
QY 244 ACCTGGTGGCTGATCATATAGTTGGCTCAGATCAGTATTCTCTAGCTCTGACCTCA--- 300
Db 78 ThrTrpProThrLeuLysProAspValCysAlaLeuAlaLaSerLeuGluSerTrp 97
QY 301 -----CCCCAGATATC----- 312
Db 98 AspileProGlyThrAspValSerSerLysArgValArgProProAspSerAspTyr 117
QY 312 ----- 312
Db 118 ThrAlaAlaLysGlnIleThrTrpGlyAlaIleGlyCysSerTyrProArgAlaArg 137
QY 313 -----CTCCATGCTCAGGATTTTATGTTGGCCAGACACCA-----AATAAT 357
Db 138 ThrArgMetAlaSerSerThrPheTyrValCysProArgAspGlyArgThrLeuSerGlu 157
QY 358 GGAACCAATTCGGAATTCCTGAGATTTCTTTGTAAACAATGGAACCTGTGTAACCTCT 417
Db 158 AlaArgArgCysGlyGlyLeuGluSerLeuTyrCysLysGluTrpAspCysGluThrThr 177
QY 418 AATGATGATATTGG-----AAATGGCCCAACC 444
Db 178 GlyThrGlyTyrTrpLeuSerLysSerSerLysAspLeuIleThrValLysTrpAspGln 197
QY 445 TCTCAGCAGGATAGGTAAGTTTCTTATGTCAACCTATACCTATACCTCTGGACATTT 504
Db 198 AsnSerGluTrpThrGlnLysPheGln-----GlnCys 208
QY 505 AATTACCTGACCTGATTAGAACTGGAAGCCCAAGTGTCTCTCTCCAGCTAGATTAC 564
Db 209 HisGlnThrGlyTrp-----CysAsnPro----- 216
QY 565 CTAAGAAATAAGTTTCACTGAGAAAGGAAAAACAGAAAATATCTCTAAATGGGTAAATGTT 624

```

```

||||| 217 LeuLysIleAspPheThrAspLysGlyLysLeuSerLys-----AspTrpIleThrGly 234
QY 625 ATGCTCTGGGAAGTATATATTATGGAGGCTCGGTTAAACAACAGGCTCCATTCTAACT 684
Db 235 LysThrTrpGlyLeuArgPheTyr---ValSerGly---HisProGlyValGlnPheThr 252
QY 685 ATTCGCTCAAAATAAACACGCTCGAGCTCCAATGGCTATAGACCAAAATACGGTCTTG 744
Db 253 IleArgLeuLysIleThrAsnMet---ProAlaValAlaValGlyProAspLeuValLeu 271
QY 745 ACGGTCAC-----AGACCCCAACCAAA 768
Db 272 ValGluGlnGlyProProArgThrSerLeuAlaLeuProProLeuProProArgGlu 291
QY 769 GGACGAGGACCATCC-----TCTAAC-----ATAACTCTGGATCAGAC 807
Db 292 AlaProProSerLeuProAspSerAsnSerThrAlaLeuAlaThrSerAlaGlnThr 311
QY 808 CCCACTGAGTCTTAACAGCAGCAGCTAAATG-----GGGGCAAAA 846
Db 312 ProThrValArgLysThrIleValThrLeuAsnThrProProThrThrGlyAspArg 331
QY 847 CTTTATAGCTCATCCAGGAGCTTTTCAAGCTCTTAAGCTTAACTCCAGACTCCAGAGGCTACC 906
Db 332 LeuPheAspLeuValGlnGlyAlaPheLeuThrLeuAsnAlaThrAsnProGlyAlaThr 351
QY 907 TCTTCTTGTGGCTGATCTTGTAGCTTCGGGCCACCTTACTATGAGGAAGTGTAGAGA 966
Db 352 GluSerCysTrpLeuCysLeuAlaMetGlyProProTyrTyrGluAlaIleAlaSerSer 371
QY 967 GGGAAATTCATATGTACAAAAGAACATAGAGACCAATGCATCGGATCCCAAAATAAG 1026
Db 372 GlyGluValAlaLysSerThrAsp---LeuAspArgCysArgTrpGlyThrGlnGlyLys 390
QY 1027 CTTACCTTACTGAGGTTTCTGGAAGAGGACCTGCATAGAGAAAGTTTCCCACTCCAC 1086
Db 391 LeuThrLeuThrGluValSerGlyHisGlyLeuCysIleGlyLysValProPheThrHis 410
QY 1087 CACACCTTGTAAACACACTGAAGCTTTAATCAAACTCTGAGACTCAATATCTGTA 1146
Db 411 GlnHisLysCysAsnGlnThrLeuSerIleAsnSerSerGlyAspHisGlnTyrLeuLeu 430
QY 1147 CCTGGTTATCAGAGTGGTGGCATGTAATCTGATTAAACCTCTGTTTCCACCTTG 1206
Db 431 ProSerAsnHisSerTrpTrpAlaCysSerThrGlyLeuThrProCysLeuSerThrSer 450
QY 1207 GTTTTAAACCAACTAAAGATTTTTCATTATGGTCCAAATGTTCCCGAGTGTTATTC 1266
Db 451 ValPheAsnGlnThrArgAspPheCysIleGlnValGlnLeuIleProArgIleTyr 470
QY 1267 TATCCCGAAAAGCAATCCTTGATGATATGATATGACTACAGAANTCATCCACAAAGAGAA 1326
Db 471 TyrProGluGluValLeuLeuGlnAlaTyrAspAsnSerHisProArgThrLysArgGlu 490
QY 1327 CCATATCTCTGACACTTGTCTGATGCTCGGACTTCGAGTGGCAGCAGGTTAGGAGAA 1386
Db 491 AlaValSerLeuThrLeuAlaValLeuLeuGlyLeuGlyIleThrAlaGlyIleGlyThr 510
QY 1387 GGAACAGCTCCCTGCTCAGGACCAACAGCAGCTAGAAAACAGACTTAGTAACCTACAT 1446
Db 511 GlySerThrAlaLeuIleLysGlyProIleAspLeuGlnGlnGlyLeuThrSerLeuGln 530
QY 1447 CGAATTGTAACAGAGATCTCCAAAGCCCTAGAAAATCTCTCAGTAACCTGGAGGATCC 1506
Db 531 IleAlaIleAspAlaAspLeuArgAlaLeuGlnAspSerValSerLysLeuGluAspSer 550
QY 1507 CTAACCTCTCTTATCTGAAGTAGTCCACAGAAAGAGAGGTTAGATTATTATTCTA 1566
Db 551 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeu 570
QY 1567 AAAGAGGAGGATTATGTGTAGCTTGAAGGAGGAATGCTGTTTTTATGTGGATCATTC 1626

```

Db 571 LysGluGlyGlyLeuCybAlaAlaLeuLysGluGlyCysPheThrIleAspHisSer 590
Qy 1627 GGGCCATCAGACTCCATGAACAGCTTAGAGAAAGTTGGAGAACGTCGAAGGAA 1686
Db 591 GlyAlaValArgAspSerMetLysLysLeuLysGluLysLeuAspLysArgGlnLeuGlu 610
Qy 1687 AAGGAAACTACTCAAGGGTGGTTTCAGGAGTGGTTTCAACAGCTCTCTTTGGTTGGTACC 1746
Db 611 ArgGlnLysSerGlnAsnTrpTyrGluGlyTrpPheAsnAsnSerProTrpPheThrThr 630
Qy 1747 CTACTTTCTGCTTTAAACAGGACCCCTTAATAGTCCTCTCTCTTACTCAGAGTTGGGCA 1806
Db 631 LeuLeuSerThrIleAlaGlyProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuGlyPro 650
Qy 1807 TGTATTATTAACAGTTAATTCCTTCAATTTAGAGAACGAATAAGTGCA 1854
Db 651 CysIleIleAsnLysLeuValGlnPheIleAsnAspArgIleSerAla 666

RESULT 10
US-09-315-127-6
; Sequence 6, Application US/09315127
; Patent No. 6448390
; GENERAL INFORMATION:
; APPLICANT: The University of Tennessee, c/o Richard Cox
; TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and
; TITLE OF INVENTION: Liposome Vectors and Use in Gene and Drug Therapy
; FILE REFERENCE: 44137-5023, U. of Tennessee
; CURRENT APPLICATION NUMBER: US/09/315,127
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SEQ ID NO. 4,
; OTHER INFORMATION: envelope protein produced by retroviral vector of
; OTHER INFORMATION: seq. id no. 3
US-09-315-127-6

Alignment Scores:
Pred. No.: 1,55e-141 Length: 667
Score: 1489.50 Matches: 314
Percent Similarity: 58.0% Conservative: 90
Best Local Similarity: 45.1% Mismatches: 171
Query Match: 42.9% Indels: 121
DB: 2 Gaps: 20

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-315-127-6 (1-667)

Qy 10 ACGTTAAACCGCGCCACCTC-----CCGATTCGGGGTGGAAAGCCGAAAGACTGAAA 63
Db 11 ThrSerAsnLeuHisLeuArgHisGlnMetSerProGlySerTyrLysArgLeu--- 29
Qy 64 ATCCCTTAAGCTTCGCCTCCATCGCGGTTCCTTACTCTGTCGTAATCAACCTCTCAGACT 123
Db 30 -----IleIleLeuLeuSerCysValPheGlyGly 39
Qy 124 AATGGTATCGCATAGGACAGACAGCTGAACCTCCCATAAACCTTATCTCACTCGGTGA 183
Db 40 GlyGlyThrSerLeu---GlnAsnLysAsnProHisGlnProMetThrLeuThrTrpGln 58
Qy 184 ATTACTGACTCCGGCACAGGTATTAAATATCAACAACACTCAAGGGGAGGCTCTTTAGGA 243
Db 59 ValLeu---SerGlnThrGlyAspValValTrpAspThrLysAlaValGlnProProTrp 77
Qy 244 ACCTGGTGGCTGATCTATACGTTTGGCTCAGATCAGTTATTCTCAGTCTGACCTCA--- 300
Db 78 ThrTrpTrpProThrLeuLysProAspValCysAlaLeuAlaAlaSerLeuGluSerTrp 97
Qy 301 -----CCCCCAGATATC----- 312

Db 98 AspIleProGlyThrAspValSerSerLysArgValArgProProAspSerAspTyr 117
Qy 312 ----- 312
Db 118 ThrAlaAlaTyrLysGlnIleThrTrpGlyAlaIleGlyCysSerTyrProArgAlaArg 137
Qy 313 -----CTCCATGCTCACGGATTTTATGTTTTCGCCAGGACCCACCA-----AATAAT 357
Db 138 ThrArgMetAlaSerSerThrPheTyrValCysProArgAspGlyArgThrLeuSerGlu 157
Qy 358 GGAATAATTCGCGAATCCAGAGATTCCTTTGTAAACAATGGAATCTGTGTAACTCT 417
Db 158 AlaArgArgCysGlyGlyLeuGluSerLeuTyrCysLysGluTrpAspCysGluThrThr 177
Qy 418 AATGATGGATATTCG-----AATGGCCCAACC 444
Db 178 GlyThrGlyTyrTrpLeuSerLysSerLysAspLeuIleThrValLysTrpAspGln 197
Qy 445 TCTCAGCAGGATAGGGTAAGTTTTCCTTATGTCAACACCTATACCAGCTCTGGCAATTT 504
Db 198 AsnSerGluTrpThrGlnLysPheGln-----GlnCys 208
Qy 505 AATTACCTGCCTGGATTAGAACTGGAAGCCCCCAAGTCTCTCTCTCAGACCTAGATTAC 564
Db 209 HisGlnThrGlyTrp-----CysAsnPro----- 216
Qy 565 CTAAAAATAGTTTCTACTGAGAAAGGAAACACGAAATATCTCTAAATGGTAAATGGT 624
Db 217 LeuLysIleAspPheThrAspLysGlyLysLeuSerLys-----AspTrpIleThrGly 234
Qy 625 ATGCTCTGGGAATGGTATATTATGGAGGCTCGGGTAAACAACACCGAGCTCTCACTTAACT 684
Db 235 LysThrTrpGlyLeuArgPheTyr---ValSerGly---HisProGlyValGlnPheThr 252
Qy 685 ATTCGCTCAAAATAAACACAGCTCGAGCTCCAAATGGCTATPAGACCAAAATACGGTCTTG 744
Db 253 IleArgLeuLysIleThrAsnMet---ProAlaValAlaValGlyProAspLeuValLeu 271
Qy 745 ACGGTCNA-----AGACCCCAACCCAA 768
Db 272 ValGluGlnGlyProProArgThrSerLeuAlaLeuProProLeuProProArgGlu 291
Qy 769 GGACCGAGCACCATCC-----TCTAAC-----ATAACTTCTGGATCAGAC 807
Db 292 AlaProProProSerLeuProAspSerAsnSerThrAlaLeuAlaThrSerAlaGlnThr 311
Qy 808 CCCACTGAGTCTAAACAGCAGCACTAAATG-----GGGCAAAA 846
Db 312 ProThrValArgLysThrIleValThrLeuAsnThrProProThrThrGlyAspArg 331
Qy 847 CTTTTTAGCTCATCCAGGAGCTTTTCAAGCTCTTAAGCTCCAGACTCCAGAGGCTACC 906
Db 332 LeuPheAspLeuValGlnGlyAlaPheLeuThrLeuAsnAlaThrAsnProGlyAlaThr 351
Qy 907 TCTTCTGCTGCTATGCTTACTGCTCGGCCCACTTACTATGAAGGATGGCTAGAGA 966
Db 352 GluSerCysThrLeuCysLeuAlaMetGlyProProTyrTyrGluAlaIleAlaSerSer 371
Qy 967 GGGAAATCAATGTCACAAAAGACATAGACCAATGACATGGGGATCCCAAAATAG 1026
Db 372 GlyGluValAlaTyrSerThrAsp---LeuAspArgCysArgTrpGlyThrGlnGlyLys 390
Qy 1027 CTTACCTTACTGAGGTTTCTGGAAGAGGACCTGTCATAGGAAGGTTTCCCATCCAC 1086
Db 391 LeuThrLeuThrGluValSerGlyHisGlyLeuCysIleGlyLysValProPheThrHis 410
Qy 1087 CAACACCTTCTCAACCACTGAAGCTTTAATCAAACTCTGAGAGTCAATATCTGTGA 1146
Db 411 GlnHisLeuCysAsnGlnThrLeuSerIleAsnSerSerGlyAspHisGlnThrLeuLeu 430
Qy 1147 CCTGTTATGACAGGTGGTGGCATGTAAATCTGGATTAAACCTTGTGTTTCCACCTTG 1206
Db 431 ProSerAsnHisSerTrpTrpAlaCysSerThrGlyLeuThrProCysLeuSerThrSer 450


```
QY 1207 GTTTTAAACCAACATAAGATTGTTTCATTATGTCCTCAATGTCCTCAATGTTCCCGAGTGTATTAC 1265
Db 451 ValPheAsnGlnThrArgAspPheCysIleGlnValGlnLeuIleProArgIleTyr 470
QY 1267 TATCCCGAAAAGCAATCTTTCATGTAATATGCTACAGAAATCATCGCAAAAGAGAA 1326
Db 471 TyrProGluGluValLeuLeuGlnAlaTyrAspAsnSerHisProArgThrLysArgGlu 490
QY 1327 CCCATATCTGACACTTGTCTGTATGCTCGGACTTGGAGTGGCAGCAGGTGTAGGAACA 1386
Db 491 AlaValSerLeuThrLeuAlaValLeuLeuGlyLeuGlyIleThrAlaGlyIleGlyThr 510
QY 1387 GGAACAGCTGCTGCTGTCACGGGACCACAGCAGCTAGAAACAGGACTTAGTAACCTACAT 1446
Db 511 GlySerThrAlaLeuIleLysGlyProIleAspLeuGlnGlnGlyLeuThrSerLeuGln 530
QY 1447 CGAATTGTAACAGAGATCTCCAGCCCTAGAAAATCTGTCACTAGTAACCTGAGGATCC 1506
Db 531 IleAlaIleAspAlaAspLeuArgAlaLeuGlnAspSerValSerLysLeuGluAspSer 550
QY 1507 CTAACCTCTTATCTGAAGTAGTCTCAGAGATAGAGAGGTTAGATTATTATTCTTA 1566
Db 551 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 570
QY 1567 AAGAAGGAGGATTATGTTAGCTTGAAGGAGGAATCTGTTTTATGTGGATCATTTCA 1626
Db 571 LysGluGlyGlyLeuCysAlaAlaLeuLysGluGluCysPheTyrIleAspHisSer 590
QY 1627 GGGGCCATCAGAGCTCATCAACAGCTTAGAGAAAGGTGGAGAGCTCGAAGGAA 1686
Db 591 GlyAlaValArgAspSerMetLysLysLeuLysGluLysLeuAspLysArgGlnLeuGlu 610
QY 1687 AAGAAATCTACTCAAGGTGCTTTGAGGATGTTCAACAGCTCTCTTTGCTGGCTACC 1746
Db 611 ArgGlnLysSerGlnAsnTyrPylGluGlyThrPheAsnAsnSerProThrPheThrThr 630
QY 1747 CTACTTCTGCTTTAAACAGGACCTTAAATAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1806
Db 631 LeuLeuSerThrIleAlaGlyProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuGlyPro 650
QY 1807 TGTATTATTACAGATTAAATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1854
Db 651 CysIleIleAsnLysLeuValGlnPheIleAsnAspArgIleSerAla 666

RESULT 11
US-09-070-630-13
; Sequence 13, Application US/09070630
; Patent No. 6750206
; GENERAL INFORMATION:
; APPLICANT: Russell, Stephen James
; APPLICANT: Morling, Frances Joanne
; APPLICANT: Fielding, Adele Kay
; APPLICANT: Cosset, Francois-Leic
; APPLICANT: Cattaneo, Roberto
; TITLE OF INVENTION: Compositions and Methods for Elimination
; FILE REFERENCE: 07039-415001
; CURRENT APPLICATION NUMBER: US/09/070,630
; CURRENT FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: PCT/GB98/00710
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 60/045,164
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: GB 9705007.4
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 667
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Plasmid sequence
US-09-070-630-13
Alignment Scores:
Pred. No.: 1,55e-141 Length: 667
Score: 1489.50 Matches: 314
Percent Similarity: 58.0% Conserved: 90
Best Local Similarity: 45.1% Mismatches: 171
Query Match: 42.9% Indels: 121
DB: 2 Gaps: 20
US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-070-630-13 (1-667)
QY 10 ACGTTAAACCGCGCCACCTC-----CCGATTCCGGGTGGAAAGCCGAAAGACTGAAA 63
Db 11 ThrSerAsnLeuHisHisLeuArgHisGlnMetSerProGlySerTrpLysArgLeu--- 29
QY 64 ATCCCTTAAAGCTTCGCTCCATCGCGTGTTCCTTACTCTGTCAATAACCTCTCAGACT 123
Db 30 -----IleIleLeuLeuSerCysValPheGlyGly 39
QY 124 AATGTTATCGCATAGGAGACGCTGAACTCCCAATAAACCTTATCTCTCACCCTGGTTA 183
Db 40 GlyGlyThrSerLeu---GlnAsnLysAsnProHisGlnProMetThrLeuThrTrpGln 58
QY 184 ATTACTGACTCCGCGCACAGGTATTAAATATCAACAACACTCAAGGGGAGGCTCTCTTAGGA 243
Db 59 ValLeu---SerGlnThrGlyAspValValTrpAspThrLysAlaValGlnProProTrp 77
QY 244 ACCTGGTGGCTGATCTATATAGTTGGCTCAGATCAGTTATTCCTAGTCTGACCTCA--- 300
Db 78 ThrTrpTrpProThrLeuLysProAspValCysAlaLeuAlaSerLeuGluSerTrp 97
QY 301 -----CCCCAGATATC----- 312
Db 98 AspIleProGlyThrAspValSerSerLysArgValArgProProAspSerAspTyr 117
QY 312 ----- 312
Db 118 ThrAlaAlaTyrLysGlnIleThrTrpGlyAlaIleGlyCysSerTyrProArgAlaArg 137
QY 313 -----CTCATGCTCAGGATTTATGTTTGGCCAGGACCAACA-----AATAAT 357
Db 138 ThrArgMetAlaSerSerThrPheTyrValCysProArgAspGlyArgThrLeuSerGlu 157
QY 358 GGAATAACATTCGGGAATCCAGAGATTTCTTTGTAAACAATGGAAGTGTGTAACCTCT 417
Db 158 AlaArgArgCysGlyGlyLeuGluSerLeuTyrCysLysGluTrpAspCysGluThrThr 177
QY 418 AATGATGATATGG-----AAATGGCCAACC 444
Db 178 GlyThrGlyTyrTrpLeuSerLysSerSerLysAspLeuIleThrValLysTrpAspGln 197
QY 445 TCTCAGCAGGATAGGTAAAGTTTCTTATGTCAACACCTATACCACTCTGGCAATTT 504
Db 198 AsnSerGluTrpThrGlnLysPheGln-----GlnCys 208
QY 505 AATTACCTGACCTGGATTAGAACCTGGAGGCCCGAGTCTCTCTTCAGACCTAGATTAC 564
Db 209 HisGlnThrGlyTrp-----CysAsnPro----- 216
QY 565 CTAATAATAAGTTTCTACTGAGAAAGAAAAACAAGAAATATCTTAAATGGTAAATGGT 624
Db 217 LeuLysIleAspPheThrAspLysGlyLysLeuSerLys-----AspTrpIleThrGly 234
QY 625 ATGCTTTGGGAATGTTATATTATGGAGGCTCGGGTAAACAACAGAGCTCCATTTCTAAT 684
Db 235 LysThrTrpGlyLeuArgPheTyr---ValSerGly---HisProGlyValGlnPheThr 252
QY 685 ATTGCGCTCAAAATAAACACCTGGAGCTCCCAATGGCTATAGGACCAATAACGCTTTG 744
Db 253 IleArgLeuLysIleThrAsnMet---ProAlaValAlaValGlyProAspLeuValLeu 271
```

```
QY 745 ACGGCTCAA-----AGACCCCAACCCAA 768
Db 272 ValGluGlnGlyProProArgThrSerLeuAlaLeuProProLeuProProArgGlu 291
QY 769 GGACAGGACCATCC-----TCTAAC-----ATAACTTCTGGATCCAGAC 807
Db 292 AlaProProProSerLeuProAspSerAsnSerThrAlaLeuAlaThrSerAlaGlnThr 311
QY 808 CCCACTGAGTCTAACAGCAGCAGCTAAATG-----GGGCAAAA 846
Db 312 ProThrValArgLysThrIleValThrLeuAsnThrProProThrThrGlyAspArg 331
QY 847 CTTTITAGCCTCATCCAGGAGCTTTTCAAGCTCTTAAGCTCTTAAGCTCCAGACTCCAGAGGCTACC 906
Db 332 LeuPheAspLeuValGlnGlyAlaPheLeuThrLeuAsnAlaThrAsnProGlyAlaThr 351
QY 907 TCTTCTCTTGGCTAGTCTGAGTTCGGGCCACCTTACTTACTGAAGGAATGGCTAGAAGA 966
Db 352 GluSerCysTrpLeuCysLeuAlaMetGlyProProTyrTyrGluAlaIleAlaSerSer 371
QY 967 GGGAAATTCATGTGACAAAAGAACATAGACCAATGCATGGGGATCCAAAATAAG 1026
Db 372 GlyGluValAlaTyrSerThrAsp---LeuAspArgCysArgTrpGlyThrGlnGlyLys 390
QY 1027 CTTACCTTACTGAGGTTCTGGAAAAGGCACCTGCATAGGAAGGTTCCCCCATCCAC 1086
Db 391 LeuThrLeuThrGluValSerGlyHisGlyLeuCysIleGlyLysValProPheThrHis 410
QY 1087 CAACACCTTTGTAAACACACATGAAGCTTTAATCAAACTCTGAGAGTCAATATCTGTA 1146
Db 411 GlnHisLeuCysAsnGlnThrLeuSerIleAsnSerSerGlyAspHisGlnTyrLeuLeu 430
QY 1147 CTGTTTATGACAGGTGGTGGCATGTAAATACTGGATTAAACCCCTGTGTTTCCACCTTG 1206
Db 431 ProSerAsnHisSerTrpTrpAlaCysSerThrGlyLeuThrProCysLeuSerThrSer 450
QY 1207 GTTTTAAACCAACTAAAGATTTTGGCATTTATGGTCCAAATGTTCCTCCGAGTGTATTAC 1266
Db 451 ValPheAsnGlnThrArgAspPheCysIleGlnValGlnLeuIleProArgIleTyrTyr 470
QY 1267 TATCCGAAAAAGCAATCCTTGATCAATATGACTACAGAAATCATCGACAAAAGAGAA 1326
Db 471 TyrProGluGluValLeuLeuGlnAlaTyrAspAsnSerHisProArgThrLysArgGlu 490
QY 1327 CCCATATCTCTGACATCTGCTGTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTAGGAACA 1386
Db 491 AlaValSerLeuThrLeuAlaValLeuLeuGlyLeuGlyIleThrAlaGlyIleGlyThr 510
QY 1387 GGAACAGCTGCCCTGTCACGGGACCACAGCGCTAGAAACAGGACTTAGTAACCTACAT 1446
Db 511 GlySerThrAlaLeuIleLysGlyProIleAspLeuGlnGlnGlyLeuThrSerLeuGln 530
QY 1447 CGAATTGTAAACAGAGATCTCCACGCCCTAGAAAATCTGTCAGTAACTCGGAGGAATCC 1506
Db 531 IleAlaIleAspAlaAspLeuArgAlaLeuGlnAspSerValSerLysLeuGluAspSer 550
QY 1507 CTAACCTCTCTATCTGAAGTAGTCTCTACAGATAGAGAGGGTTAGATTTATTATTCTA 1566
Db 551 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 570
QY 1567 AAAGAAGAGGAGTTATGTAGTACCTTGAGGAGGATGCTGTTTATGTGGATCAATCA 1626
Db 571 LysGluGlyGlyLeuCysAlaAlaLeuLysGluGluCysCysPheTyrIleAspHisSer 590
QY 1627 GGGGCGCATCAGAGCTCCATGAACAAGCTTTAGAGAAAGGTTGGAGAAGCGTCGAGGAGAA 1686
Db 591 GlyAlaValArgAspSerMetLysLysLeuLysGluLysLeuAspLysArgGlnLeuGlu 610
QY 1687 AAGAAACTACTCAAGGGTGGTTTCAGGGATGGTTCAACAGGTCTCTTTGGTGGCTACC 1746
Db 611 ArgGlnLysSerGlnAsnTrpTyrGluGlyTyrPheAsnAsnSerProTrpPheThrThr 630
QY 1747 CTACTTCTGCTTTAACAGGACCCCTTAATAGTCCCTCCTCTTACTCACAGTGGGCCA 1806
```

```
Db 631 LeuLeuSerThrIleAlaGlyProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuGlyPro 650
QY 1807 TGTATTATTAAACAGTTAAATTGCTTTCATTAGAGAACGAATAAGTGCA 1854
Db 651 CysIleIleAsnLysLeuValGlnPheIleAsnAspArgIleSerAla 666

RESULT 12
US-09-075-272-5
; Sequence 5, Application US/09075272
; Patent No. 6136598
; GENERAL INFORMATION:
; APPLICANT: MILLER, A. DUSTY
; APPLICANT: WOLGAMOT, GREG
; APPLICANT: BONHAM, LYNN
; TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL
; TITLE OF INVENTION: PACKAGING CELL LINES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,272
; FILING DATE: 08-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,140
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: 14538A-003710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 673 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-075-272-5
```

```
Alignment Scores:
Pred. No.: 2,56e-140 Length: 673
Score: 1477.50 Matches: 327
Percent Similarity: 59.8% Conservative: 78
Best Local Similarity: 47.6% Mismatches: 161
Query Match: 42.6% Indels: 121
DB: 2 Gaps: 26
```

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-075-272-5 (1-673)

```
QY 70 TTAGCTTGGCTCCTCATCGCTGGTTCCTTACTCTGTCAATAACCTCTCAGACTAATGCT 129
Db 18 LeuSerPheAlaCysValAlaGly-----AlaProSerIleThr----- 30
QY 130 ATGCCATAGACAGACAGCGCTCAACTCCCATAAACCTTATCTCTCACCTGGTTAATTACT 189
Db 31 LeuAspLeuGlyAsnHis---AsnProHisAlaProValGlnGlnSerIrp----- 46
QY 190 GACTCCGGCACAGAGTATTATATATCAACAACTCAAGGGAGGCT----- 234
Db 47 -----GluValLeuAsnGluLysGlyAspValValTrpValAlaThr 60
```

QY 235 -----CCTTAGGAACCTGGTGGCTGATCTATAC-----GTTTGC---CTCAGA 276
Db 61 AlaValHisProProTrrPThrTrpProAspLeuThrProAspIleCysLysLeuAla 80
QY 277 TCAGTTATTCTCTAGT-----CTGACCTCACCCCCAGAT--- 309
Db 81 AlaGlySerProAsnTrpAspLeuProAspHisThrAspLeuAsnAsnProProSerGlu 100
QY 309 ----- 309
Db 101 GlnLysCysValProAsnGlyValGlySerThrThrGlyCysSerGlyGlnPheTyrArg 120
QY 310 ---ATCTCCATGCTCAGGATTTTATGTTGGCCAGGACACCAATAATAGGAAAA--- 363
Db 121 AlaAsnLeuArgAlaAlaGlnPheTyrValCysProGlyGlnGlyGlnLysGlyLysLeu 140
QY 364 -----CATTCGCGAAATCCAGAGATTCTTTTGTAAACAATGGAACCTGTGTAACCTCT 417
Db 141 GlnGlnGluCysArgGlyAlaSerAspTyrPheCysGlyLysTrpThrCysGluThrThr 160
QY 418 AATGATGGATATTGAAATGCCAACCTCTCAGCAGGATAGGTAAGTTTCTTATGTC 477
Db 161 GlyGluAlaTyrTrpLys---ProSerAlaAspTrpAspLeuIleThrValLysArgGly 179
QY 478 AACACCTAT-----ACGAGCTCTGGCAATTTAATACCTGACCTGGATPAGAACTGGA 531
Db 180 SerGlyTyrAspLysProAsnGlnGlyGluArgAsnProTyrLysTyrLeuAspSerGly 199
QY 532 -----AGCCCCAAG-----TGCTCTCCCTCA 552
Db 200 CysAlaLeuLysAsnTyrSerProProGlyProCysLysGlyLysTyrCysAsnPro--- 218
QY 553 GACCTAGATTACCTAAAAAATAGTTTCTCAGAGAAAGGAAACAGAAATAATCTTAAAA 612
Db 219 -----LeuLeuIleLysPheThrGluLysGlyLysGlnAlaArg---LeuSer 233
QY 613 TGGGTAATGATGTCTTGGGGAATGGTATATTATGGAGCTCGGGTAACACACAGGC 672
Db 234 TrpLeuLysGlyAsnArgTrpGlyTrpArgValTyrIleProIle---ArgAspProGly 252
QY 673 TCACATTTAATATTCGCTCAAAATAAACCCAGCTGGAGCTCAATGGCTATAGGACCA 732
Db 253 PheIlePheThrIleArgLeuThrValArgAspLeu---AlaValThrSerIleGlyPro 271
QY 733 AATACGGCTTTGACGGGTCAA----- 753
Db 272 AsnLysValLeuThrGluGlnAlaProProValAlaProAlaProProArgValProAla 291
QY 754 -----AGACCCCAACCCCAAGGACCA-----GGACCATCTCTAACAATACTTCT 798
Db 292 ValProAlaProProThrSerArgProTyrThrValGlyProSerLeuGluThrThrLeu 311
QY 799 GGATCAGACCCC-----ACTGAGTCTAACGACGACTAAATATGGGGGCAAACTT 849
Db 312 AlaSerProProLeuLeuAspThrGluAsn-----ArgLeu 323
QY 850 TTTAGCTCTACAGGAGCTTTTCAAGCTCTTAACCTCCAGCTCCAGAGGCTACCTCT 909
Db 324 ValSerLeuValGlnGlyAlaPheLeuValLeuAsnArgThrAsnProAsnMetThrGln 343
QY 910 TCTTGTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGAAATGGCTGAAGAGGG 969
Db 344 SerCysTrpLeuCysTyrAlaSerAsnProProTyrTyrGluGlyIleAlaGlnThrArg 363
QY 970 AATTCATGTGCAAAAGAACATAGAGACCAATGCATGGGATCCCAAAATAAGCTT 1029
Db 364 ThrTyrAsnIleThrSerAspHis---SerGlnCysLeuTrpGlyGluAsnArgLysLeu 382
QY 1030 ACCCTTACTGAGTTCTGGAAGGACCTGCATAGAAAGGTTCCCTCCATCCACCAA 1089
Db 383 ThrLeuThrAlaValSerGlyAsnGlyLeuCysLeuGlyGlnValProGlnAspLysTrp 402

QY 1090 CACCTTTGTAAACACACTGAAGCCCTTTAATCAAACTCTGAGAGTCAATATCTGTACCT 1149
Db 403 HisLeuCysAsnGlnThrGlnAsnIleArgProAsnLysGlyGlnTyrLeuValPro 422
QY 1150 GGTATTACAGAGTGGTGGCATGTAATACATCGATTAACCCCTGTGTTCCACCTTGTT 1209
Db 423 ProIleAspThrValTrpAlaCysAsnThrGlyLeuThrProCysIleSerMetSerVal 442
QY 1210 TTTAACCAAACTAAAGATTTTTCATATGTCCTCAAAATTTGTTCCCGAGTGTATTACTAT 1269
Db 443 PheAsnSerSerLysAspPheCysIleLeuValGlnLeuIleProArgLeuLeuTyrHis 462
QY 1270 CCCGAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCAGCAAAAGAGAGAACCC 1329
Db 463 AspAspSerSerPheLeuAspLysPheGluHisArg---ValArgTrpLysArgGluPro 481
QY 1330 ATATCTCTGACACTGTCTGATGCTCGACTCGAGTGGAGTG---CGACGAGGTGAGGACAA 1386
Db 482 IleThrLeuThrLeuAlaValLeuLeuGlyLeuGlyValAlaAlaIleGlyValGlyThr 501
QY 1387 GGAACAGCTGCTCTGCTGACGGGACACAGCAGCTAGAAAACAGGACTTAGTAACCTACAT 1446
Db 502 GlyThrAlaAlaLeuIleGlnThrProArgTyrPheGlu-----GluLeuArg 517
QY 1447 CGAATTGTAAACAGAGATCTCAAGCCCTAGAAAATCTGTACAGTAACCTGGAGAAATCC 1506
Db 518 ThrAlaMetAspThrAspLeuArgAlaIleGluHisSerIleThrLysLeuGluGluSer 537
QY 1507 CTAACCTCCTTATCTGAAGTAGTCTCTACAGAAATAGAGAGGTTAGATTATTATTCTTA 1566
Db 538 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 557
QY 1567 AAAGAACGAGGATTTATGTAGCTTCAAGGAGGAATGCTGTTTTTTATGTGATCATTTCA 1626
Db 558 LysGluGlyGlyLeuCysAlaAlaLeuLysGluGluCysCysPheTyrValAspHisSer 577
QY 1627 GGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAAAAGTTGGAGAGCGTCGAGAGGAA 1686
Db 578 GlyValIleLysAspSerMetAlaLysLeuArgGluArgLeuAspIleArgGlnArgGlu 597
QY 1687 AAGAAACTACTCAAGGCTGTTTGGAGGATGTTTCAACAGGCTCTCTTTGGTTGGCTACC 1746
Db 598 ArgGluSerLysGlnGlyTrpPheGluSerTrpPheAsnLysSerProTrpLeuThrThr 617
QY 1747 CTACTTTCTGCTTTAACGAGCACCTTAATAGTCTCTCTCTGTTTACTCACAGTTGGGCCA 1806
Db 618 LeuLeuSerThrIleAlaGlyProLeuIleLeuLeuLeuLeuLeuLeuThrPheGlyPro 637
QY 1807 TGTATTATTAAAGTTAATTGCTTTCATTTAGAGAACGAATAAGTGCAGTCCAGATCATG 1866
Db 638 CysIleLeuAsnLysLeuValAlaPheIleArgGluArgIleAsnAlaValGlnValMet 657
QY 1867 GTACTTAGACAACAGTACCAA 1887
Db 658 ValLeuLysGlnGlnTyrGln 664
RESULT 13
; Sequence 14, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 665
; TYPE: PRT

ORGANISM: Moloney murine leukemia virus
FEATURE:
OTHER INFORMATION: env protein
US-09-309-572-14

Alignment Scores:

Pred. No.: 2,756-124 Length: 665
Score: 1319.00 Matches: 291
Percent Similarity: 55.7% Conservative: 86
Best Local Similarity: 43.0% Mismatches: 198
Query Match: 38.0% Indels: 102
DB: 2 Gaps: 20

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-309-572-14 (1-665)

```
QY 49 CCGAAAGACGTAAGTCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGPCA 108
    |||||
DB 16 ProArgGlyProLeuIleProLeuIleLeuMetLeuArgGlyValSerThrAlaSer 35
QY 109 ATAACTCTCAGACTAATGGTATGGCATAGGACACGCTGAACTCCATAAACCCCTTA 168
    |||||
DB 36 ProGlySer-----SerProHisGlnValTyr 44
QY 169 TCTCTACCTGTTAAATTAAGTCTCGGCTCCATCGCGTGGTTCCTTACTCTGPCA 228
    |||||
DB 45 AsnIleThrTrpGluValThr----AsnGlyAspArgGluThrValTrpAlaThrSerGly 63
QY 229 GAGGCTCTTTAGGAACCTGGTGGCTGATCTATAC-----GTTGGCTC----- 273
    |||||
DB 64 AsnHisProLeuTrpThrTrpTrpProAspLeuThrProAspLeuCysMetLeuAlaHis 83
QY 274 -----AGATCAGTATTTCCTAGTCTGACCTCACCC 303
    |||||
DB 84 HisGlyProSerTyrTrpGlyLeuGluTrpGlnSerProPheSerProGlyPro 103
QY 304 CCA----- 306
DB 104 ProCysCysSerGlyGlySerSerProGlyCysSerArgAspCysGluGluProLeuThr 123
QY 307 -----GATATCCTCCAT 318
DB 124 SerLeuThrProArgCysAsnThrAlaTrpAsnArgLeuLysLeuAspGlnThrThrHis 143
QY 319 GCT-----CAGGATTTTATGTTGGCCAGGACCA-----CCAAATAATGGAACAT 366
    |||||
DB 144 LysSerAsnGluGlyPheTyrValCysProGlyProHisArgProArgGluSerLysSer 163
QY 367 TCGGGAATCCAGAGATTCTTTGTAAACAATGGAACCTGTGTAACCTCTTAATGATGGA 426
    |||||
DB 164 CysGlyGlyProAspSerPheTyrCysAlaTyrTrpGlyCysGluThrThrGlyArgAla 183
QY 427 TATTGGAATGGCCAACTCTCAGCAGGATAGGTAAGTTTCTTATGTCAACACCTAT 486
    |||||
DB 184 TyrTrpLys---ProSerSerSerTrpAsp-----PheIleThrValAsnAsnAsn 199
QY 487 ACCAGCTCTGACAAATTTAATTACCTGACCTGGATAGGACGCGCCCAAGTGTCT 546
    |||||
DB 200 LeuThrSerAspGln-----AlaValGlnValCysLysAspAsnLysTrpCysAsn 216
QY 547 CTTCTCAGACCTAGATTACCTAAATAAGTTTCACTGAGAAAGGAAACAAAGAAATATC 606
    |||||
DB 217 Pro-----LeuValIleArgPheThrAspAlaGlyArgArg-----Val 229
QY 607 CTAATAATGGTAAATGGTATGTTTGGGGAATGGTATATTATGGAGGCTCGGTAACAA 666
    |||||
DB 230 ThrSerTrpThrThrGlyHisTyrTrpGlyLeuArgLeuTyr---ValSerGlyGlnAsp 248
QY 667 CGAGCTCCATCTAATCTATTGCGCTCAAAATAAACACGAGCTGGAGCCTCCAAATGGCTATA 726
    |||||
DB 249 ProGlyLeuThrPheGlyIleArgLeuArgTyrGlnAsnLeuGlyProArgValProIle 268
QY 727 GGACCAAAATACGCTCTCAGCGGTCAAAGACCC-----CCAAACCAAGGACCAAGA 777
    |||||
```

```
DB 269 GlyProAsnProValLeuAlaAspGlnProLeuSerLysProLysProValLysSer 288
QY 778 CCATCCTCTACATAACTTCTGGATCAGACCCCTGAGTCTTAACAGACACACTAAATG 837
    |||||
DB 289 ProSerValThrLysProProSerGlyThrProLeuSerProThrGlnLeuProProAla 308
QY 838 GGGGCA-----AAACTTTTGTAGCTCATCCAGGAGGCTTTTCAAGCTCTTAACTCCACG 891
    |||||
DB 309 GlyThrGluAsnArgLeuLeuAsnLeuValAspGlyAlaTyrGlnAlaLeuAsnLeuThr 328
QY 892 ACTCAGAGGCTACCTCTTCTGTGTGCTATGCTTAGCTTCGGGCCACCTTACTATGAA 951
    |||||
DB 329 SerProAspLysThrGlnGluCysTrpLeuValAlaGlyProProTyrTrpGlu 348
QY 952 GGAATGGCTGACAGAGGGAATTCATGTCAGAAAAGAACATAGAGACCATGACATGG 1011
    |||||
DB 349 GlyValAlaValLeuGlyThrTyrSerAsnHisThrSerAlaProAlaAsnCysSerVal 368
QY 1012 GGATCCCAAAATAAGCTTACCTTACTGAGGTTCTTGGAAAAGGACCTGCATAGGAAAG 1071
    |||||
DB 369 AlaSerGlnHisLysLeuThrLeuSerGluValThrGlyGlnGlyLeuCysIleGlyAla 388
QY 1072 GTTCCCCCATCCCAACACCTTTGTAAACCACTGAAGCCTTTAATCAAACTCTGAG 1131
    |||||
DB 389 ValProLysThrHisGlnAlaLeuCysAsnThrThr-----GlnThrSerSer 404
QY 1132 -----AGTCAATATCTGGTACCTGGTATGACAGGTGGTGGGATGTAATACTGGATTA 1185
    |||||
DB 405 ArgGlySerTyrTyrLeuValAlaProThrGlyThrMetTrpAlaCysSerThrGlyLeu 424
QY 1186 ACCCTTGTGTTCCTCCACTGCTGTTTAAACCAAACTAAAGATTTTGCATTATGTCCTCAA 1245
    |||||
DB 425 ThrProCysIleSerThrThrIleLeuAsnLeuThrThrAspTyrCysValLeuValGlu 444
QY 1246 ATTGTTCCCCGAGTGTATTACTATCCCGAAAAGCAATCTTGTATGATATGACTACAGA 1305
    |||||
DB 445 LeuTrpProArgValThrTyrHisSerProSerTyrValTyrGlyLeuPheGlu---Arg 463
QY 1306 AATCATCGCAAAAGAGAGAACCCATATCTCTGACACTTGTCTGTGATGCTCGGA----- 1359
    |||||
DB 464 SerAsnArgHisLysArgGluProValSerLeuThrLeuAlaLeuLeuLeuGlyGlyLeu 483
QY 1360 -----CTTGAGTGGCAGCAGGTGTAGGAACAGCAACAGCTGCGCTGGTCAACGGACCA 1413
    |||||
DB 484 ThrMetGlyGlyIleAlaAlaGlyIleGlyThrGlyThrAlaLeuMetAlaThrGln 503
QY 1414 CAGCAGCTAGAAAACAGGACTTAGTAACCTACATCGAATTTGTAACAGAGATCTCCAAAGCC 1473
    |||||
DB 504 Gln-----PheGlnGlnLeuGlnAlaValGlnAspAspLeuArgGlu 518
QY 1474 CTAGAAAATCTGTCAGTAACCTGGAGAAATCCCTAACCTCTTATCTGAAGTAGTCTTA 1533
    |||||
DB 519 ValGluLysSerIleSerAsnLeuGluLysSerLeuThrSerLeuSerGluValValLeu 538
QY 1534 CAGAAATAGAGAGGTGTAGATTATTTCTTAAAGAGGAGGATTTGTGTAGCCTTG 1593
    |||||
DB 539 GlnAsnArgArgGlyLeuAspLeuLeuPheLeuLysGluGlyGlyLeuCysAlaAlaLeu 558
QY 1594 AAGGAGGAATCTGTTTTTATGTGATCATTCAGGGCCATCAGAGACTCCATGAACAG 1653
    |||||
DB 559 LysGluGluCysCysPheTyrAlaAspHisThrGlyLeuValArgAspSerMetAlaLys 578
QY 1654 CTTAGAGAAAGGTTGGAGACGCTCGAAGGAGGAAAGAACTACTCAAGGCTGGTTGAG 1713
    |||||
DB 579 LeuArgGluArgLeuAsnGlnArgGlnLysLeuPheGluSerThrGlnGlyTrpPheGlu 598
QY 1714 GGATGGTTCAACAGTCTCTTTGTGCTGCTACCTCTTCTGCTTTTAAACAGGACCCCTTA 1773
    |||||
DB 599 GlyLeuPheAsnArgSerProTrpPheThrThrLeuIleSerThrIleMetGlyProLeu 618
QY 1774 ATAGTCTCTCTCTGTTACTCAGAGTTGGGCGCATGTATTATTAAAGAGTTAAATTCGCTTC 1833
    |||||
DB 619 IleValLeuLeuMetIleLeuLeuPheGlyProCysIleLeuAsnArgLeuValGlnPhe 638
```


Db 1052 ArgGlySerTyrTyrLeuValAlaProThrGlyThrMetTyrAlaCysSerThrGlyLeu 1071
Qy 1186 ACCCTTGTGTTTCCACTTGTTTAAACCAACTAAAGATTTTTCATTATGTCCTCAA 1245
Db 1072 ThrProCysIleSerThrThrIleLeuAsnLeuThrThrAspTyrCysValLeuValGlu 1091
Qy 1246 ATTGTTCCCGAGTGATTACTATCCGAAAAGCAATCCTTGATGATATGACTACAGA 1305
Db 1092 LeuTyrProArgValThrTyrHisSerProSerTyrValTyrGlyLeuPheGlu---Arg 1110
Qy 1306 AATCATCGACAAAGAGAGAACCCATATCTCTGACACTTGCTGTGATGCTCGGA----- 1359
Db 1111 SerAsnArgHisLysArgGluProValSerLeuThrLeuAlaLeuLeuLeuGlyGlyLeu 1130
Qy 1360 -----CTTGAGTGGCAGCGGTGTAGGAACAGGAACAGCTGCCCTGGTCACGGGACCA 1413
Db 1131 ThrMetGlyIleAlaAlaGlyIleGlyThrGlyThrThrAlaLeuMetAlaThrGln 1150
Qy 1414 CAGCAGCTAGAAACAGGACTTAGTAACCTACATCGAATTGTAAACAGAAGATCTCCAGCC 1473
Db 1151 Gln-----PheGlnGlnLeuGlnAlaAlaValGlnAspSerLeuArgGlu 1165
Qy 1474 CTAGAAAATCTGTGAGTAACCTGAGGAATCCCTAACCTCTTATCTGAAGTAGTCCTA 1533
Db 1166 ValGluLysSerIleSerAsnLeuGluLysSerLeuThrSerLeuSerGluValValLeu 1185
Qy 1534 CAGAAATAGAGAGGTTAGATTATTTATTTCTTAAAGAGGAGGATTATGTAGCCTTG 1593
Db 1186 GlnAsnArgArgGlyLeuAspLeuPheLeuLysGluGlyLeuCysAlaAlaLeu 1205
Qy 1594 AAGAGGAAATGCTGTTTATGTGATCAATTCAGGGGCCATCAGAGCTCCATGAACAAG 1653
Db 1206 LysGluGluCysCysPheTyrAlaAspHisThrGlyLeuValArgAspSerMetAlaLys 1225
Qy 1654 CTTAGAGAAAGTTGGAGAGCGTCGAGGGAAGAAAGAACTACTCAAGGGTGGTTTCAG 1713
Db 1226 LeuArgGluArgLeuAsnGlnArgLysLeuPheGluSerThrGlnGlyTyrPheGlu 1245
Qy 1714 GGATGGTTCAACAGGTCTCTTTGGTTGGCTACCTTCTTCTTAAACAGGACCCCTTA 1773
Db 1246 GlyLeuPheAsnArgSerProTyrPheThrThrLeuIleSerThrIleMetGlyProLeu 1265
Qy 1774 ATAGTCCTCTCTGTTACTCACAGTTGGGCCATGTATTATTAACAAGTTAATTGCCTTC 1833
Db 1266 IleValLeuLeuMetIleLeuLeuPheGlyProCysIleLeuAsnArgLeuValGlnPhe 1285
Qy 1834 ATTAGAGACGAATAAGTCAGATCCAGATCATGTGCTACTTAGACAACAGTAC 1884
Db 1286 ValLysAspArgIleSerValValGlnAlaLeuValLeuThrThrGlnGlnTyr 1302

Search completed: February 14, 2006, 16:18:41

Job time : 98.4616 secs

This Page Blank (uspto)

Db 11 ThrSerAsnLeuHisLeuArgHisGlnMetSerProGlySerTrpIysArgLeu--- 29
Qy 64 ATCCCTTAAGCTTCGCCTCCATCCGCTGGTTCCTTACTCTCTCAATAACCTTCAGACT 123
Db 30 -----llelleLeuLeuSerCysValPheGlyGly 39
Qy 124 AATGGTATGCGATAGGAGACAGCCTGAACCTCCCAATAAACCTTATCTCTCACCTGGGTA 183
Db 40 GlyGlyThrSerLeu---GlnAsnLysAsnProHisGlnProMetThrLeuThrTrpGln 58
Qy 184 ATTACTGACTCGGCACAGGTATTATATCAACAACACTCAAGGGAGGCTCTTTAGGA 243
Db 59 ValLeu---SerGlnThrGlyAspValValTrpAspThrLysAlaValGlnProProTrp 77
Qy 244 ACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTATTCCTAGTCTGACCTCA--- 300
Db 78 ThrTrpTrpProThrLeuLysProAspValCysAlaLeuAlaLeuAlaSerLeuGluSerTrp 97
Qy 301 -----CCCCCAGATATC----- 312
Db 98 AspileProGlyThrAspValSerSerSerLysArgValArgProProPronSpSerAspTyr 117
Qy 312 ----- 312
Db 118 ThrAlaAlaTyLysGlnlethrTrpGlyAlaIleGlyCysSerTyrProArgAlaArg 137
Qy 313 -----CTCCATGCTCAGGATTTTATGTTTGCCTCCAGGACCCCA-----AATAAT 357
Db 138 ThrArgMetAlaSerSerThrPheTyrValCysProArgAspGlyArgThrLeuSerGlu 157
Qy 358 GGAAACACATGGGAAATCCAGAGATTTCTTTGTAAACAATGAACCTGTGTAACTCT 417
Db 158 AlaArgArgCysGlyGlyLeuGluSerLeuTyrCysLysGluTrpAspCysGluThrThr 177
Qy 418 AATGATGATATG-----AAATGGCCAAACC 444
Db 178 GlyThrGlyTyrTrpLeuSerLysSerSerLysAspLeuLeuThrValLysTrpAspGln 197
Qy 445 TCTCAGCAGGATAGGTAAAGTTTCTTATGTGTACACACTATACAGCTCTGGACAATT 504
Db 198 AsnSerGluTrpThrGlnLysPheGln-----CysAsnPro----- 208
Qy 505 AATTACCTGACCTGGATTAGAACTGGAAGCCCAAGTCTCTCTCTTACAGCTCTAGATTAC 564
Db 209 HisGlnThrGlyTrp----- 216
Qy 565 CTAATAAATAGTTTCACCTCAGAAAGGAAACAAAGAAATATCTCTAAATGGTAAATGGT 624
Db 217 LeuLysIleAspPheThrAspLysGlyLysLeuSerLys-----AspTrpIleThrGly 234
Qy 625 ATGCTCTGGGNAATGGTATATTATGGAGGCTCGGGTAAACAAACAGGCTCCATTCTAACT 684
Db 235 LysThrTrpGlyLeuArgPheTyr---ValSerGly---HisProGlyValGlnPheThr 252
Qy 685 ATTCCCTCAAAATAAACAGCTGAGGCTCCCAATGGCTATAGGACCAAAATACGCTCTTG 744
Db 253 lleArgLeuLysIleThrAsnMet---ProAlaValAlaValGlyProAspLeuValLeu 271
Qy 745 ACGGCTCAA-----AGACCCCAACCCCAA 768
Db 272 ValGluGlnGlyProProArgThrSerLeuAlaLeuProProProProArgGlu 291
Qy 769 GGACAGGACCAATCC-----TCTAAC-----ATACTTCTGGATCAGAC 807
Db 292 AlaProProProSerLeuProAspSerAsnSerThrAlaLeuAlaThrSerAlaGlnThr 311
Qy 808 CCCACTGAGCTAAACAGCAGCAGCTAAATG-----GGGCAAAA 846
Db 312 ProThrValArgLysThrIleValThrLeuAsnThrProProProThrThrGlyAspArg 331
Qy 847 CTTTTAGCCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAGGCTACC 906
Db 332 LeuPheAspLeuValGlnGlyAlaPheLeuThrLeuAsnAlaThrAsnProGlyAlaThr 351

Qy 907 TCTCTTGTGCTATGCTTAGCTTCGGGCCCCACCTTACTATGAGGAATGCTAGAGA 966
Db 352 GluSerCysTrpLeuCysLeuAlaMetGlyProProTyrTyrGluAlaIleAlaSerSer 371
Qy 967 GGGAAATCAATGTCACAAAGAAACATAGAGACCAATGACATGGGATCCCAAATTAAG 1026
Db 372 GlyGluValAlaTySerThrAsp---LeuAspArgCysArgTrpGlyThrGlnGlyLys 390
Qy 1027 CTTACCTTACTAGGTTTCTGGAAAGGACCTGTCATAGGAAGGTTCCCCATCCAC 1086
Db 391 LeuThrLeuThrGluValSerGlyHisGlyLeuCysIleGlyLysValProPheThrHis 410
Qy 1087 CAACACCTTTCTAACACACAGCCCTTAACTCAAACTCTGAGAGTCAATATCTGGTA 1146
Db 411 GlnHisLeuCysAsnGlnThrLeuSerIleAsnSerSerGlyAspHisGlnTrpLeuLeu 430
Qy 1147 CCTGTTATGACAGGTGGTGGCATGTAATACTGGAATTAACCCCTGTGTTCACCTTG 1206
Db 431 ProSerAsnHisSerTrpTrpAlaCysSerThrGlyLeuThrProCysLeuSerThrSer 450
Qy 1207 GTTTTAAACCAACTAAAGATTTTGCATTTATGTCCTCAAAATGTTCCCGAGTGTATTAC 1266
Db 451 ValPheAsnGlnThrArgAspPheCysIleGlnValGlnLeuIleProArgIleTyr 470
Qy 1267 TATCCGAAAAAGCAATCTTGTGATGATATGACTACAGAAATCATCGACAAAGAGAGAA 1326
Db 471 TyrProGluGluValLeuLeuGlnAlaTyAspAsnSerHisProArgThrLysArgGlu 490
Qy 1327 CCCATATCTCGACACTTGTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTAGAAACA 1386
Db 491 AlaValSerLeuThrLeuAlaValLeuGluGlyLeuGlyIleThrAlaGlyIleGlyThr 510
Qy 1387 GGAACAGCTGCCCTGGTCCCGGACCCAGCAGCTAGAGAAACAGGACTTAGTAACCTACAT 1446
Db 511 GlySerThrAlaLeuIleLysGlyProIleAspLeuGlnGlnGlyLeuThrSerLeuGln 530
Qy 1447 CGAATTGTAAACAGAGATCTCAAGCCCTGAGAAATCTGTCAAGTAACCTGGAGGAATCC 1506
Db 531 IleAlaIleAspAlaAspLeuArgAlaLeuGlnAspSerValSerLysLeuGluAspSer 550
Qy 1507 CTAACCTCTTATCTGAAGTAGTCTCAGAAATAGAGAGGTTAGATTTATTATTCTA 1566
Db 551 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 570
Qy 1567 AAAGAGAGAGATTATGTTAGCTTGAAGAGGAAATGCTGTTTATGTCGATCATCA 1626
Db 571 LysGluGlyGlyLeuCysAlaAlaLeuLysGluCysCysPheTyrIleAspHisSer 590
Qy 1627 GGGGCCATCAGAGACTCCATCAAGCTTAGAGAAAGGTTGGAGAACGCTCGAAGGAA 1686
Db 591 GlyAlaValArgAspSerMetLysLysLeuLysGluLysLeuAspLysArgGlnLeuGlu 610
Qy 1687 AAGAAACTACTCAAGGCTGTTGAGGATGGTTCAACAGGTCTCTTTGTTGGCTACC 1746
Db 611 ArgGlnLysSerGlnAsnTrpTyrGluGlyTrpPheAsnAsnSerProTrpPheThrThr 630
Qy 1747 CTACTTTCTGCTTTAAACAGGACCCCTTAATAGTCTCTCTCTGTTACTCAGCTGGGCA 1806
Db 631 LeuLeuSerThrIleAlaGlyProLeuLeuLeuLeuLeuLeuLeuIleLeuGlyPro 650
Qy 1807 TGTATTATTAAACAGTTAAATTCCTTATTAGAGAACCAATAAGTGCA 1854
Db 651 CysIleIleAsnLysLeuValGlnPheIleAsnAspArgIleSerAla 666

RESULT 2

S70395

env polyprotein - Friend murine leukemia virus (strain FB29)

N:Contains: knob protein gp6; R protein; spike protein p15E

C:Species: Friend murine leukemia virus

A:Variety: strain FB29

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S70395

[illegible]

Db 466 TyrValTyrSerGlnPheGlu---LysSerHisArgHisLysArgGluProValSerLeu 484
QY 1339 ACACCTGCTGTGATGCTCGGA-----CTTGAGTGGCAGCAGAGGTGTAGGAACA 1386
Db 485 ThrLeuAlaLeuLeuGlyGlyLeuThrMetGlyGlyIleAlaAlaGlyValGlyThr 504
QY 1387 GGAACAGCTGCCCTGCTACGGGACCACAGCAGCTAGAAACAGAGCTTAGTAACCTACAT 1446
Db 505 GlyThrThrAlaLeuValAlaThrGlnGln-----PheGlnGlnLeuHis 519
QY 1447 CGAATTGTGAACAGAGATCTCCAGCCCTAGAAAATCTGTCAGTAACTCGAGGAAATCC 1506
Db 520 AlaAlaValGlnAspAspLeuGlyGluValGlnLysSerIleThrAsnGluLysSer 539
QY 1507 CTAACCTCTTATCTGAAGTAGTCTCTACAGAAATAGAGAGGTTAGATTATTATTCTA 1566
Db 540 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeu 559
QY 1567 AAAGAACGAGGATATGTGTAGCTTGAAGGAGGAATGCTGTTTTATGTGGATCATCA 1626
Db 560 LysGluGlyGlyLeuCysAlaAlaLeuLysGluCysPheTyrAlaAspHisThr 579
QY 1627 GGGGCCATCAGACCTCCATCAACAAAGCTTAGAGAAAGTTGGAGAGCGTCGAAGGGA 1686
Db 580 GlyLeuValArgAspSerMetAlaLysLeuArgGluArgLeuSerGlnArgGlnLysLeu 599
QY 1687 AAGAAACTACTCAAGGCTGTTGAGGGATGGTTCAACAGGCTCTCTTTGGTGGCTACC 1746
Db 600 PheGluSerSerGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrPheThrThr 619
QY 1747 CTACTTCTGCTTTAAACAGACCCCTTAATAGTCTCTCTCTGTTACTCAGAGTTGGGCCA 1806
Db 620 LeuLeuSerThrIleMetGlyProLeuIleLeuLeuLeuLeuLeuLeuPheGlyPro 639
QY 1807 TGATTATTAAACAGTAATGTCCTCATTAGAGAACGAATAAGTCAGATCCAGATCATG 1866
Db 640 CysIleLeuAsnArgLeuValGlnPheValLysAspArgIleSerValValGlnAlaLeu 659
QY 1867 GTACTTAGACAACAGTAC 1884
Db 660 ValLeuThrGlnGlnTyr 665
RESULT 4
VCMPV
env polyprotein precursor - Friend murine leukemia virus (strain PVC-211)
N:Alternate names: coat polyprotein
N:Contains: knob protein gp76; R protein; spike protein p15E
C:Species: Friend murine leukemia virus
C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C:Accession: A38210; S35476
R:Masuda, M.; Remington, M.P.; Hoffman, P.M.; Ruscetti, S.K.
J. Virol. 66, 2798-2806, 1992
A:Title: Molecular characterization of a neuropathogenic and nonerythroleukemogenic vari
A:Reference number: A38210; MUID:92219364; PMID:1560524
A:Accession: A38210
A:Molecule type: DNA
A:Residues: 1-676 <SPE>
A:Cross-references: UNIPROT:P26803; UNIPARC:UPI000012A03D; GB:M931134; NID:g331898; PIDN:
R:Remington, M.P.; Hoffman, P.M.; Ruscetti, S.K.; Masuda, M.
Nucleic Acids Res. 20, 3249, 1992
A:Title: Complete nucleotide sequence of a neuropathogenic variant of Friend murine leuk
A:Reference number: S35474; MUID:92319660; PMID:1620621
A:Accession: S35476
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-676 <REM>
A:Cross-references: UNIPARC:UPI000012A03D; EMBL:M931134; NID:g331898; PIDN:AAA46478.1; PT
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protei
F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-676/Product: env polyprotein #status predicted <ENV>
F:36-619/Domain: extracellular #status predicted <EXT>
F:36-479/Product: knob protein gp76 #status predicted <KGP>
F:476-479/Region: cleavage processing #status predicted
F:480-659/Product: spike protein p15E #status predicted <SP>
F:486-502/Region: hydrophobic #status predicted
F:620-636/Domain: transmembrane #status predicted <TM>
F:637-676/Domain: intracellular #status predicted <INT>
F:660-676/Product: R protein #status predicted <RPT>
F:46, 202, 336, 368, 408, 444/Binding site: carbohydrate (Asn) (covalent) #status predicted
Alignment Scores:
Pred.:No.: 2.88e-99 Length: 676
Score: 1345.00 Matches: 297
Percent Similarity: 55.6% Conservative: 86
Best Local Similarity: 43.1% Mismatches: 190
Query Match: 38.8% Indels: 116
DB: 1 Gaps: 22
US-10-723-552-3_COPY_5620_7533 (1-1914) x VCMVPV (1-676)
QY 49 CCGAAAAGACTGAAATCCCTTAAGCTTCGCCTCCATCGCGTGTCTTACTCTGTCA 108
Db 16 ProArgAspLeuLeuLeuProLeuLeu-----PheLeuSerLeu--- 29
QY 109 ATAACCTCTCAGACTAATGGTATGCGC---ATAGGAGACAGCCTGAACTCCCATAAACCC 165
Db 30 -----LysGlyAlaArgSerAlaAlaProGlySerSerProHisGlnVal 44
QY 166 TTATCTCTACCTGGTTAATTAATCTGACTCCGGCAGAGTATTATATCAACAACACTCAA 225
Db 45 TyrAsnIleThrTrpGluValThr---AsnGlyAspArgGluThrValTrpAlaIleSer 63
QY 226 GGGGAGCTCTTTAGGAACCTGGTGGCTGATCTATAC-----GTTTGCCTC--- 273
Db 64 GlyAsnHisProLeuTrpThrTrpTrpProAspLeuThrProAspLeuCysMetLeuAla 83
QY 274 -----AGATCAGTTATTCTCTGCTGCTGACCTCA 300
Db 84 LeuSerGlyProProHisTrpGlyLeuGluTyrArgAlaProTyrSerSerProGly 103
QY 301 CCCCCA----- 306
Db 104 ProProCysCysSerGlySerSerGlyAsnArgAlaGlyCysAlaArgAspCysAspGlu 123
QY 307 -----GAT 309
Db 124 ProLeuThrSerLeuThrProArgCysAsnThrAlaTrpAsnArgLeuLysLeuAspGln 143
QY 310 ATCTCTCCAT-----GCTCAGCGATTATGTTTGGCCAGACCA-----CCAAATAAT 357
Db 144 ValThrHisLysSerSerGlyGlyPheTyrValCysProGlySerHisArgProArgLys 163
QY 358 GGAACAATTCGGGAATCCAGAGATTTCTTTTGTAAACAATGGAAGTGTGTAACCTCT 417
Db 164 AlaLysSerCysGlyGlyProAspSerPheTyrCysAlaSerTrpGlyCysGluThrThr 183
QY 418 AATGATGATATGGAATGGCAATGCGCACTCTCAGCAGATAGGTAAGTTTCTTATGTC 477
Db 184 GlyArgAlaTyrTrpLys---ProSerSerSerTrpAspTyrIleThrValAsp----- 200
QY 478 AACACCTATACACAGCTCTCGACAAATTAATTACCTGACCTGGATTAGAACTGGAGCCCC 537
Db 201 AsnAsnLeuThrThr-----AsnGlnAlaGlnValCysLysAspAsnLys 216
QY 538 AAGTGTCTCTTCAGACCTAGATTACCTAAAAATAAGTTTCACTGAGAAGGAAAAACA 597
Db 217 TrpCysAsnPro-----LeuAlaIleGlnPheThrAsnAlaGlyLysGln 231
QY 598 GAAAATATCTAAATGGTAAATGGTATGCTCTTGGGAATGGTATATATTGAGGCTCG 657
Db 232 -----ValThrSerTrpThrIleGlyHisTyrTrpGlyLeuArgLeuTyr---ValSer 248

QY 658 GGTAAACACACAGGCTCCATTCTAACTATTGGCTCAAAATAAACCAGCTGGAGCTCCA 717
Db : : : : :
QY 249 GlyGlnAspProGlyLeuThrPheGlyIleArgLeuLysTyrGlnAsnLeuGlyProArg 268
Db : : : : :
QY 718 ATGGCTATAGGACCAAAATACGGTCTTGACGGGTCAAGACCC---CCAAACCAAGGACCA 774
Db : : : : :
QY 269 ValProIleGlyProAsnProValLeuAlaAspGlnLeuSerPheProLeuProAsnPro 288
Db : : : : :
QY 775 GGACCATCTCTAACATAACTTCTGGATCAGACCCACCTGAGTCTAACAGCAGCACTAAA 834
Db : : : : :
QY 289 LeuProLys-----ProAlaLysSerProSerAlaSerAsnSerThrProThr 304
QY 835 ATG-----GGGCGCAAAACTT 849
Db : : : : :
QY 305 LeuIleSerProSerProAlaProThrGlnProProProAlaGlyThrGlyAspArgLeu 324
QY 850 TTTAGCCTCATCCAGGGAGCTTTTCAAGCTCTTAACCTCCACGACTCCAGGCTACCTCT 909
Db : : : : :
QY 325 LeuAsnLeuValGlnGlyAlaTyrGlnAlaLeuAsnLeuThrAsnProAspLysThrGln 344
QY 910 TCTGTGCTGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGNATGGCTAGAGAGGG 969
Db : : : : :
QY 345 GluCysTrpLeuCysLeuValSerAlaProProTyrTyrGluGlyValAlaValLeuGly 364
QY 970 AAATTCAATGTGACAAAAGACATAGACCAATGCACATGGGATCCCAAAATAAGCTT 1029
Db : : : : :
QY 365 ThrTyrSerAsnHisThrSerAlaProAlaAsnCysSerAlaGlySerGlnHisLysLeu 384
QY 1030 ACCCTTACTAGGTTTCTGGAAAAGGCACCTGCATPAGGAAAGGTTTCCCCCATCCACCAA 1089
Db : : : : :
QY 385 ThrLeuSerGluValThrGlyGlnGlyLeuCysIleGlyThrValProLysThrHisGln 404
QY 1090 CACCTTTGTACCACACTGAAGCCTTTAATCAACCTCTGAGTCAATATCTGGTACT 1149
Db : : : : :
QY 405 AlaLeuCysAsnThrThr-----LeuLysThrGlyLysGlySerTyrThrLeuValAla 422
QY 1150 GGTATGACAGGTGGTGGCATGTAATACTGATTAAACCCCTTGTTGGTTCCACCTTGTT 1209
Db : : : : :
QY 423 ProAlaGlyThrMetTrpAlaCysAsnThrGlyLeuThrProCysLeuSerAlaThrVal 442
QY 1210 TTTAACCAAACTAAAGATTTTGCATTATGTTGCCAAATGTTTCCCGGAGTGTATTACTAT 1269
Db : : : : :
QY 443 LeuAsnArgThrThrAspTyrCysValLeuValGluLeuTyrProArgValThrTyrHis 462
QY 1270 CCGGAAAAGCAATCTTGATGAATATGACTACAGAAATCATCGACAAAGAGGAGAACCC 1329
Db : : : : :
QY 463 ProProSerTyrValTyrSerGlnPheGlu---LysSerTyrArgHisLysArgGluPro 481
QY 1330 ATATCTCTGACACTGTGCTGATGCTCCGA-----CTTGGAGTGGCAGCAGGT 1377
Db : : : : :
QY 482 ValSerLeuThrLeuAlaLeuLeuLeuLeuGlyGlyLeuThrMetGlyGlyIleAlaAlaGly 501
QY 1378 GTAGGAACAGGAACAGCTGCCCTGTGTCACGGGACACAGCAGCTAGAACAGGACTTAGT 1437
Db : : : : :
QY 502 ValGlyThrGlyThrThrAlaLeuValAlaThrGlnGln-----PheGln 516
QY 1438 AACCTAGATCGAATGTTGAACAGAGATCTCCAGCCCTAGAAAAATCTGTCAGTAACCTG 1497
Db : : : : :
QY 517 GlnLeuHisAlaAlaValGlnAspAspLeuLysGluValGluLysSerIleThrAsnLeu 536
QY 1498 GAGGAATCCCTAACTCTTATCTCAAGTAGTCTACAGATAGAGAAGGAGGTTAGATTTA 1557
Db : : : : :
QY 537 GluLysSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 556
QY 1558 TTATTTCTAAAGAGGAGGATTTATGTTAGCTTTGAAGGAGGAATGCTGTTTTTATGTG 1617
Db : : : : :
QY 557 LeuPheLeuLysGluGlyGlyLeuCysAlaAlaLeuLysGluLysCysPheTyrAla 576
QY 1618 GATCATTCAGGGCCATCAGACACTCCATGAACAGCTTAGAGAAAGGTTGAGNAGGCT 1677
Db : : : : :
QY 577 AspHisThrGlyLeuValArgAspSerMetAlaLysLeuArgGluLeuThrGlnArg 596
QY 1678 CGAAGGAAAAGGAACTACTCAAGGGTGGTTTGAGGGATGGTTCAACAGGCTCTCTTGG 1737

Db 597 GlnLysLeuPheGluSerSerGlnGlyTrpPheGluGlyLeuPheAsnArgSerProTrp 616
QY 1738 TTGGCTACCTTCTGCTTTTAAACAGGACCCCTTAATAGTCCCTCTCTTACTACA 1797
Db 617 PheThrThrLeuIleSerThrIleMetGlyProLeuIleIleLeuLeuLeuIleLeu 636
QY 1798 GTTGGCCATGTATTATTAAACAAGTTAATGCCCTTCTATTAGAGAACGAATAAGTGCAGTC 1857
Db 637 PheGlyProCysIleLeuAsnArgLeuValGlnPheValLysAspArgIleSerValVal 656
QY 1858 CAGATCATGGTACTTACACACAGTAC 1884
Db 657 GlnAlaLeuValLeuThrGlnGlnTyr 665
RESULT 5
VCMVKA
env polyprotein precursor - radiation murine leukemia virus (strain Kaplan)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp76; coat protein p15E; R protein
C;Species: radiation murine leukemia virus
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: B42743
R;Poliquin, L.; Bergeron, D.; Fortier, J.L.; Paquette, Y.; Bergeron, R.; Rassart, E.
J. Virol. 66, 5141-5146, 1992
A;Title: Determinants of thymotropism in Kaplan radiation leukemia virus and nucleotide
A;Reference number: A42743; MUID:92333703; PMID:1629969
A;Accession: B42743
A;Molecule type: DNA
A;Residues: 1-665 <POL>
C;Cross-references: UNIPROT:P31794; UNIPARC:UPI000012A042; GB:I93052; NID:G332065; PIDN:
C;Genetics:
A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protei
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-665/Product: env polyprotein #status predicted <ENV>
F;32-604/Domain: extracellular #status predicted <EXT>
F;32-467/Product: coat protein gp76 #status predicted <GP1>
F;464-467/Region: cleavage processing #status predicted
F;468-644/Product: coat protein p15E #status predicted
F;474-490/Region: hydrophobic #status predicted
F;605-621/Domain: transmembrane #status predicted <TM1>
F;622-665/Domain: intracellular #status predicted <INT>
F;645-665/Product: R protein #status predicted <RPP>
F;43,199,211,324,356,396,400,432/Binding site: carbohydrate (Asn) (covalent) #status pre
Alignment Scores:
Pred. No.: 3.46e-99 Length: 665
Score: 1344.00 Matches: 297
Percent Similarity: 58.2% Conservative: 92
Best Local Similarity: 44.8% Mismatches: 179
Query Match: 38.8% Indels: 100
DB: 1 Gaps: 21
US-10-723-552-3_COPY_5620_7533 (1-1914) x VCMVKA (1-665)
QY 97 CTTACTCTGTCATTAACCTCTCAGACTAATGGTATGCCATAGGAGACAGCCTGAACCTCC 156
Db : : : : :
QY 21 IleValLeuLeuIleLeuGlyArgValAsnProValAlaLeuLeuGlyAsnSer-----Pro 38
QY 157 CATAAACCTTATCTCTCACCTGGTTAATTACTACTCCTCCGGCACAGGATTAAATATCAAC 216
Db : : : : :
QY 39 HisGlnValPheAsnLeuSerTrpGluValThrAsnGluAspArgGlu---ThrValTrp 57
QY 217 AACACTCAAGGGAGGCTCTTTTAGAACCTGGTGGCTGATCTATAC-----GTTTGC 270
Db : : : : :
QY 58 AlaIleThrGlyAsnHisProLeuTrpThrTrpProAspLeuThrProAspLeuCys 77
QY 271 CTCAGATCAGTTATT---CCTAGTCTG-----ACCTCACCC 303
Db : : : : :
QY 78 MetLeuAlaLeuHisGlyProSerTyrTrpGlyLeuGluTyrGlnAlaProPheSerPro 97

QY 304 CCA----- 306
Db 98 ProProGlyProProCysSerArgSerGlySerThrProGlyCysSerArgAsp 117
QY 306 ----- 306
Db 118 CysGluGluProLeuThrSerTyrThrProArgCysAsnThrAlaTrpAsnArgLeuLys 137
QY 307 -----GATATCTCCATGCTCAC-----GGATTTTATGTTTGCACAGGACCA----- 348
Db 138 LeuSerLysValThrHisAlaHisAsnGluGlyPheTyrValCysProGlyProHisArg 157
QY 349 CCAATATATGGAACATTCGGAAATCCAGAGATTCTTTTGTAAACAATGGAACGT 408
Db 158 ProArgTrpAlaArgSerCysGlyGlyProGluSerPheTyrCysAlaSerTrpGlyCys 177
QY 409 GTAACCTCTAATGATGATATTTGAAATGGCAACCTCTCAGCAGGATAGGTAAGTTT 468
Db 178 GluThrThrGlyArgAlaSerTrpLys---ProSerSerSerTrpAspTyrIleThr--- 195
QY 469 TCTTATGTCAACACCTATACCAGCTCTCGACAATTTAATTACCTGCACCTGGATTAGAACT 528
Db 196 -----ValSerAsnAsnLeuThrSerGlyGln----- 204
QY 529 GGAAGCCCCAAGTGTCTCTCTTCAGAC---CTAGATTACCTAAATAAATGTTTCACTGAG 585
Db 205 AlaThrProValCysLysAsnAsnThrTrpCysAsnSerLeuThrIleArgPheThrSer 224
QY 586 AAAGGAAACAAAGAAATATCTAAATGGCTAAATGGTATGTCTTGGGAAATGGTATAT 645
Db 225 LeuGlyLysGlnAlaThr-----SerTrpValThrGlyHisTrpTrpGlyLeuArgLeu 242
QY 646 TATGGAGGCTCGGTAAACAAACAGGCTCCATTTCTAACTATTCGCTCCAAATAAACCAG 705
Db 243 Tyr---ValSerGlyHisAspProGlyLeuIlePheGlyIleArgLeuLysIleThrAsp 261
QY 706 CTGAGAGCTCAANTGGCTATAGGACCAATAGGTCTTTGAGGGTCAAGACCCCAAC 765
Db 262 SerGlyProArgValProIleGlyProAsnProValLeuSerAspGlnArgProProSer 281
QY 766 CAAGGACCAGGACCATCTCTAAACATAACTTCTGGATCAGACCCACCTGAGTCTAACAGC 825
Db 282 Gln-----ProArgSerProProHisSerAsnSerThrProThrGluThrProLeu 298
QY 826 ACG-----ACTAAATGGGGCAAAACTTTTATAGCTTCATCCAGGGA 867
Db 299 ThrLeuProGluProProProAlaGlyValGluAsnArgLeuLeuAsnLeuValLysGly 318
QY 868 GCTTTTCAAGCTCTTAATCTCACGACTCCAGAGGCTACCTCTTCTTGTGGCTATGCTTA 927
Db 319 AlaTyrGlnAlaLeuAsnLeuThrSerProAspArgThrGlnGluCysTrpLeuCysLeu 338
QY 928 GCTTCGGGGCCACCTTACTATGAGGAATGCTAGAAGAGGAATTCATATGTCACAAA 987
Db 339 ValSerGlyProProTyrTyrGluGlyValAlaValLeuGluGlyThrTyrSerAsnHisThr 358
QY 988 GAACATAGAGACCACTGACATGGGATCCCAATAAGCTTACCTTACTGAGTTTCT 1047
Db 359 SerAlaProAlaAsnCysSerValAlaSerGlnHisLysLeuThrLeuSerGluValThr 378
QY 1048 GGAAGAGCCACCTGCATAGGAAAGTTCCCCCAATCCCAACACCTTTTGTAAACCACT 1107
Db 379 GlyArgGlyLeuCysValGlyAlaValProLysThrHisGlnAlaLeuCysAsnThrThr 398
QY 1108 GAAGCCTTTAATCAAACTCTGAGAGTCATATCTGGTACTCTGTTATGACAGGTGTGG 1167
Db 399 Gln-----AsnThrSerGlyGlySerTyrTyrLeuAlaAlaProAlaGlyThrIleTrp 416
QY 1168 GCATGTATATCTGATTAACCCCTTGTGTTTCCACCTTGGTTTAAACCAAACTTAAGAT 1227
Db 417 AlaCysAsnThrGlyLeuThrProCysLeuSerThrThrValLeuAsnLeuThrThrAsp 436
QY 1228 TTTTGCATTATGGTCCAAATTTGTTCCCGAGTGATTTACTATATCCGAAAAAGCAATCCTT 1287

Db 437 TyrCysValLeuValGluLeuTrpProArgValThrTyrHisSerProSerTyrValTyr 456
QY 1288 GATCAATATGACTACAGAAATCATCGACAAAGAGAGAACCATATCTCTGACACTTCT 1347
Db 457 HisGlnPheGluGlyArg---AlaLysTyrLysArgGluProValSerLeuThrLeuAla 475
QY 1348 GTGATGCTCGGA-----CTTGGAGTGCAGCAGGTAGGAAACAGGAAACAGCT 1395
Db 476 LeuLeuLeuGlyGlyLeuThrMetGlyGlyIleAlaAlaGlyValGlyThrGlyThrThr 495
QY 1396 GCCCTGCTCAGGGACCAACAGCAGCTAGAAACAGAGCTTAGTAACCTACATCGAATTGTA 1455
Db 496 AlaLeuVal---AlaThrGlnGlnLeuGlnAlaAla-----Val 507
QY 1456 ACAGAAAGTCTCCAAAGCCCTAGAAAAATCTGTCAAGTAACCTGGAGGAATCCCTAACCTCC 1515
Db 508 HisAspAspLeuLysGluValGluLysSerIleThrAsnLeuGluLysSerLeuThrSer 527
QY 1516 TTATCTGAAGTAGTCTTACAGAAATAGAGAGGGTTAGATTATTATTCTTAAAAAGAGGA 1575
Db 528 LeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeuPheLeuLysGluGly 547
QY 1576 GGATATGTGTAGCTTGAAGAGGAATGCTGTTTTTATGTGATCATTCAGGGCCATC 1635
Db 548 GlyLeuCysAlaAlaLeuLysGluGluCysCysPheTyrAlaAspHisThrGlyValVal 567
QY 1636 AGAGACTCCATGACAAAGCTTAGAAGAGCTTGGAGAAGCGTCAAGGGAAAGAGAACT 1695
Db 568 ArgAspSerMetAlaLysLeuArgGluArgLeuAsnGlnArgLysLeuPheGluSer 587
QY 1696 ACTCAAGGGTGTGAGGGATGGTCAACAGAGTCTCTTTGGTTGGCTACCTACTTTCT 1755
Db 588 GlyGlnGlyTrpPheGluArgLeuPheAsnGlySerProTrpPheThrThrLeuIleSer 607
QY 1756 GCTTTAACAGACCTTAATAGTCTCTCTCTGTACTCACAGTTGGGCCATGATATTATT 1815
Db 608 ThrIleMetGlyProLeuIleValLeuLeuLeuIleLeuLeuGlyProCysIleLeu 627
QY 1816 AACAAAGTTAATGCTTTCATTAGAGAAGCAATAGTGCAGTCACAGATCATGTTACTTAGA 1875
Db 628 AsnArgLeuValGlnPheValLysAspArgIleSerValValGlnAlaLeuValLeuThr 647
QY 1876 CAACAGTACCAGAAAGCCGCTCTAGC 1899
Db 648 GlnGlnTyrHisGlnLeuLysSer 655
RESULT 6
VCMVVR
env polyprotein precursor - radiation murine leukemia virus
N;Alternate names: coat polyprotein
N;Contains: coat protein gp76; coat protein p15E; R protein
C;Species: radiation murine leukemia virus
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: C26183
R;Merregaert, J.; Janowski, M.; Reddy, E.P.
Virology 158, 88-102, 1987
A;Title: Nucleotide sequence of a radiation leukemia virus genome.
A;Reference number: A94362; MUID: 87207680; PMID: 3033897
A;Molecule type: DNA
A;Accession: C26183
A;Residues: 1-665 <MER>
A;Cross-references: UNIPROT:P11268; UNIPARC:UPI000012A041; GB:K03363; GB:M18449; NID:G3
C;Genetics:
A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane prote
F;1-31/Domain: signal sequence #status predicted <SIG>
F;32-665/Product: env polyprotein #status predicted <ENV>
F;32-604/Domain: extracellular #status predicted <EXT>
F;32-467/Product: coat protein gp76 #status predicted <GP1>
F;464-467/Region: cleavage processing #status predicted
F;468-644/Product: coat protein p15E #status predicted <GP2>

F;474-490/Region: hydrophobic #status predicted
F;605-621/Domain: transmembrane #status predicted <TM>
F;622-665/Domain: intracellular #status predicted <INT>
F;645-665/Product: R protein #status predicted <RP>
F;43,199,211,324,356,396,400,432/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 8,67e-99 Length: 665
Score: 1339.00 Matches: 295
Percent Similarity: 58.1% Conservative: 93
Best Local Similarity: 44.2% Mismatches: 180
Query Match: 38.6% Indels: 100
DB: 1 Gaps: 21

US-10-723-552-3_COPY_5620_7533 (1-1914) x VCMVVR (1-665)

Qy	97	CTTACTCTGTCNATAA	CTCTCAGAC	TAATGGTATGCGCATAGGACAGACGCTGNACTCC	156		
		:::	:::	::: ::: :::			
Db	21	IleValLeuLeuIleLeuGlyArgValAsnProValAlaLeuGlyAsnSer	-----	Pro	38		
Qy	157	CATAAAACCC	TTATCTCTC	CACCTGGTAACTTACTGACCTCCGCCACAGGTATTAATATCAAC	216		
		::: ::: ::: :::		:::			
Db	39	HisGlnValPheAsnLeuSerT	TrpGluValThrAsnGluAspArgGlu	---ThrValT	57		
Qy	217	AACACTCAAGGGAGGCTCTTTAGGAACCTGTGGCTGATCTATAC	-----	GT	270		
		:::		:::			
Db	58	AlaIleThrGlyAsnHisProLeuT	TrpT	TrpProAspLeuThrProAspLeuCys	77		
Qy	271	CTCAGATCAGTTAT	---CCTAGTCTG	-----	303		
		::: ::::		:::			
Db	78	MetLeuAlaLeuHisGlyProSerTyrT	TrpGlyLeuGluTyrGlnAlaProPheSerPro	97			
Qy	304	CCA	-----	306			
		::					
Db	98	ProProGlyProProCysCysSerGlySerSerGlySerThrProGlyCysSerArgAsp	117				
Qy	306	-----	306				
Db	118	CysGluGluProLeuThrSerTyrThrProArgCysAsnThrAlaT	AsnArgLeuLys	137			
Qy	307	-----GATATCCTCCATGCTCAC	-----	348			
		:::		:::			
Db	138	LeuSerLysValThrHisAlaHisAsnGluGlyPheTyrValCysProGlyProHisArg	157				
Qy	349	CCAAATAATGGAAACATTCGGGAAATCCAGACAGATTCTTTGTAAACAATGGAACGT	408				
		::: ::: ::: :::					
Db	158	ProArgT	TrpAlaArgSerCysGlyGlyProGluSerPheTyrCysAlaSerT	TrpGlyCys	177		
Qy	409	GTAACCTCTAATGATGGATATTTGAAATGCGCAACCTCTCAGCAGGATAGGGTAAGTTT	468				
		::: ::: ::: :::		:::			
Db	178	GluThrThrGlyArgAlaSerT	TrpLys	---ProSerSerSerT	TrpAspTyrIleThr	195	
Qy	469	TCTTATGTCAACACTATACAGCTCTGGACAATTAATTACCTGACCTGGATTAGAACT	528				
		::: ::: ::: :::					
Db	196	-----ValSerAsnAsnLeuThrSerGlyGln	-----	204			
Qy	529	GGAGCCCCCAAGTCTCTCTCTTCAGAC	---CTAGATTACCTAAAAATAAGTTTCACTGAG	585			
		::: ::: ::: :::		:::			
Db	205	AlaThrProValCysLysLysAsnAsnThrT	TrpCysAsnSerLeuThrIleArgPheThrSer	224			
Qy	586	AAAGGAAAAACAAGAAAAATATCTAAAAATGGGTAATGGTATGTCTTGGGGAATGGTATAT	645				
		::: ::: ::: :::		:::			
Db	225	LeuGlyLysGlnAlaThr	-----SerT	TrpValThrGlyHisT	TrpT	TrpGlyLeuArgLeu	242
Qy	646	TATGGAGGCTCGGGTAAACAACACAGGCTCCATTTCTAATTCGCCTCAAAAATAAACACAG	705				
		::: ::: ::: :::					
Db	243	Tyr	---ValSerGlyHisAspProGlyLeuIlePheGlyIleArgLeuLysIleThrAsp	261			
Qy	706	CTGGAGCTCCAATGGCTATAGGACCAATAACGTCTTGCGGTCAAGACCCCAACC	765				
		::: ::: ::: :::					
Db	262	SerGlyProArgValProIleGlyProAsnProValLeuSerAspGlnArgProProSer	281				
Qy	766	CAAGGACGAGGACCATCTCTTAAACATAACTTCTCGATCAGACACCCCACTGAGTCTTAACAGC	825				

Db 628 AsnArgLeuValGlnPheValIysAspArgIleSerValValGlnAlaLeuValLeuThr 647
 QY 1876 CAACAGTACCAAGCCCTCTAGC 1899
 Db 648 GlnGlnTyrHisGlnLeuLysSer 655
 RESULT 7
 VCWEEK
 N: polyprotein - AKV murine leukemia virus
 N: Contains: knob protein gp76; R protein; spike protein p15E
 C: Species: AKV murine leukemia virus
 C: Date: 05-Apr-1983 #sequence revision 03-Aug-1984 #text_change 09-Jul-2004
 C: Accession: A92995; A93448; A03984
 R: Herr, W.
 J. Virol. 49, 471-478, 1984
 A: Title: Nucleotide sequence of AKV murine leukemia virus.
 A: Reference number: A92995; MUID: 84115072; PMID: 6319746
 A: Accession: A92995
 A: Molecule type: genomic RNA
 A: Residues: 1-669 <HER>
 A: Cross-references: UNIPROT: P03386; UNIPARC: UPI000012A039; GB: J01998; GB: J01999; GB: K000
 R: Herr, W.; Corbin, V.; Gilbert, W.
 Nucleic Acids Res. 10, 6931-6944, 1982
 A: Title: Nucleotide sequence of the 3' half of AKV.
 A: Reference number: A93448; MUID: 83090450; PMID: 6294621
 A: Accession: A93448
 A: Molecule type: DNA
 A: Residues: 1-34, 'R', 36-462, 'K', 464-591, 'K', 593-669 <HE2>
 A: Cross-references: UNIPARC: UPI0000174A49
 C: Genetics:
 A: Gene: env
 C: Superfamily: type C retrovirus env polyprotein
 C: Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protein
 F: 1-31/Domain: signal sequence #status predicted <SIG>
 F: 32-470/Product: knob protein gp76 #status predicted <KNS>
 F: 471-650/Product: spike protein p15E #status predicted <SPK>
 F: 651-669/Product: R protein #status predicted <RPT>
 F: 43,199,327,359,399/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
 Pred. No.: 8.67e-99 Length: 669
 Score: 1339.00 Matches: 295
 Percent Similarity: 58.1% Conservative: 94
 Best Local Similarity: 44.1% Mismatches: 174
 Query Match: 38.6% Indels: 106
 DB: 1 Gaps: 22

US-10-723-552-3_COPY_5620_7533 (1-1914) x VCWEEK (1-669)
 QY 97 CTACTCTGTCAATAACCTCTCAGACTAATGGTATGGCATAGGAGACAGCTGAACCTC 156
 Db 21 IleValLeuLeuIleLeuGlyGlyValAsnProValThrLeuGlyAsnSer-----Pro 38
 QY 157 CATAAACCTTATCTCAGCTGTTAACTAGCTACCTCGGACAGAGTATATATCAAC 216
 Db 39 HisGlnValPheAsnLeuThrTrpGluValThr---AsnGlyAspArgGluThrValTrp 57
 QY 217 AACACTCAAGGGAGGCTCTTTAGGAACCTGTGGCTGATCTATAC-----GTTTGC 270
 Db 58 AlaIleThrGlyAsnHisProLeuThrTrpTrpProAspLeuThrProAspLeuCys 77
 QY 271 CTCAGATCAGTATTT---CTAGTCTG-----ACCTCACCC 303
 Db 78 MetLeuAlaLeuHisGlyProSerTyrTrpGlyLeuGluTyrArgAlaProPheSerPro 97
 QY 304 CCA----- 306
 Db 98 ProProGlyProProCysCysSerGlySerSerAspSerThrProGlyCysSerArgAsp 117
 QY 306 ----- 306
 Db 118 CysGluGluProLeuThrSerTyrThrProArgCysAsnThrAlaTrpAsnArgLeuLys 137

QY 307 -----GATATCCCTCCATGCTCAC-----GGATTTTATGTTTGCCTCAGGACCA----- 348
 Db 138 LeuSerLysValThrHisAlaHisAsnGlyGlyPheTyrValCysProGlyProHisArg 157
 QY 349 CCAATATATGGAACAACATTCGGAAATCCAGAGATTTCTTTTGTAAACATATGGAACGT 408
 Db 158 ProArgTrpAlaArgSerCysGlyGlyProGluSerPheTyrCysAlaSerTrpGlyCys 177
 QY 409 GTAACCTCTAATGATGATATTGAAATGGCAACCTCTCAGCAGGATAGGTAGTATTT 468
 Db 178 GluThrThrGlyArgAlaSerTrpLys---ProSerSerSerTrpAspTyrIleThrVal 196
 QY 469 TCTTATGTCAACACTATACAGCTCTGGCAATTTAATTACCTGACCTGGATTAGAACT 528
 Db 197 Ser-----AsnAsnLeuThrSerAsp-----Gln 204
 QY 529 GGAAGCCCCAAGTGCTCTCTCTTCAGAC-----CTAGATTACCTAAATAAAGTTTCACTGAG 585
 Db 205 AlaThrProValCysLysGlyAsnGluTrpCysAsnSerLeuThrIleArgPheThrSer 224
 QY 586 AAGGAAAAACAAGAAAAATATCTAAAAATGGTAAATGGTATGCTCTGGGGAATGCTATAT 645
 Db 225 PheGlyLysGlnAlaThr-----SerTrpValThrGlyHisTrpTrpGlyLeuArgLeu 242
 QY 646 TAGGAGGCTCGGTAAACAACAGGCTCCATTCTAACTATTCGCCTCAAAATAAACCCAG 705
 Db 243 Tyr---ValSerGlyHisAspProGlyLeuIlePheGlyIleArgLeuLysIleThrAsp 261
 QY 706 CTGGAGGCTCCATGGCTATAGGACCAATACGGTCTTTCGGGTCAAAGACCC----- 759
 Db 262 SerGlyProArgValProIleGlyProAsnProValLeuSerAspArgArgProProSer 281
 QY 760 -----CAAACCCAGGACAGGACCACTCTCTAACATACTTCTGGATCAGACCCC 810
 Db 282 ArgProArgProThrArgSerProProProSerAsn-----SerThrPro 296
 QY 811 ACTGAGTCTAACAGCAGC-----ACTAAATGGGGGCAAAACTTTT 852
 Db 297 ThrGluThrProLeuThrLeuProGluProProAlaGlyValGluAsnArgLeuLeu 316
 QY 853 AGCTCTATCCAGGAGGCTTTTCAAGCTCTTAACCTCAGCAGCTCAGAGGCTACCTCTCT 912
 Db 317 AsnLeuValLysGlyAlaTyrGlnAlaLeuAsnLeuThrSerProAspLysThrGlnGlu 336
 QY 913 TGTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAGAGGAAA 972
 Db 337 CysTrpLeuCysLeuValSerGlyProProTyrTrpGlyValAlaValLeuGlyThr 356
 QY 973 TTCAATGTGACAAAGAACATAGAGACCAATGCACATGGGGATCCCAAAATAAGCTTACC 1032
 Db 357 TyrSerAsnHisThrSerAlaProAlaAsnCysSerValAlaSerGlnHisLysLeuThr 376
 QY 1033 CTTACTGAGGTTTCTGGAAGGACCTGCATAGGAAGGTTCCCTCCATCCCAACAC 1092
 Db 377 LeuSerGluValThrGlyGlnGlyLeuCysIleGlyAlaValProLysThrHisGlnVal 396
 QY 1093 CTTTGTAAACCACTGAGGCTTTTAATCAACCTCTGAG---AGTCAATATCTGGTACCT 1149
 Db 397 LeuCysAsnThrThrGln-----LysThrSerAspGlySerTyrTyrLeuAlaAla 413
 QY 1150 GGTATGACAGGTTGGTGGCATGTAATCTGGAATTAACCCCTTGTGTTTCCACCTCGGT 1209
 Db 414 ProThrGlyThrThrTrpAlaCysSerThrGlyLeuThrProCysIleSerThrThrIle 433
 QY 1210 TTTAACCAACTAAAGATTTTGCATATGGTCAAAATGTTTCCCGAGGTATTACTAT 1269
 Db 434 LeuAspLeuThrThrAspTyrCysValLeuValGluLeuTrpProArgValThrThrHis 453
 QY 1270 CCGGAAAAACAATCCTTGATGATATGACTACAGAAATCATCGACAAAGAGAGACCC 1329
 Db 454 SerProSerTyrValTyrHisGlnPheGlu---ArgArgAlaLysTyrLysArgGluPro 472
 QY 1330 ATATCTCTGACACTTGTGTGATGCTCGGA-----CTTGAGTGGCAGCAGGT 1377

Db SerGluValThrGlyGlnGlyLeuCysIleGlyAlaValProLysThrHisGlnValLeu 397
Qy TGTAAACACATGAAGCCTTTAATCAAACTCTGAG---AGTCAATATCTGGTACCTGGT 1152
Db CysAenThrThrGln-----LysThrSerAspGlySerTyrTyrLeuValAlaPro 414
Qy TATGACAGGTGGTGGCATGTAATACCTGATTAACCCCTTGCTGTTTCCACCTGGTTTTT 1212
Db ThrGlyThrThrTrpAlaCysSerThrGlyLeuThrProCysIleSerThrThrIleLeu 434
Qy AACCAAACTAAAGATTTTTCATATGTCCTCAATTTGTCCTCCGAGTGTATTACTATCCC 1272
Db AsnLeuThrThrAspTyrCysValLeuValGluLeuTrpProArgValThrTrpHisSer 454
Qy GAAAAAGCAATCTTGATGAATATGACTACAGAAATCATCGAAAAAGAGAGAACCCATA 1332
Db ProSerTyrValTyrHisGlnPheGlu--ArgArgAlaLysTyrLysArgGluProVal 473
Qy TCTCTGACACTGCTGTCATCTCTCGA-----CTTGAGTGGGAGCGGTGTA 1380
Db SerLeuThrLeuAlaLeuLeuGlyGlyLeuThrMetGlyGlyIleAlaAlaGlyVal 493
Qy GGAACAGGAACAGCTGCCTCGTCACGGGACCACAGCAGCTAGAAAAACAGGACTTAGTAAC 1440
Db GlyThrGlyThrThrAlaLeuValAlaThrGlnGln-----PheGlnGln 508
Qy CTACATCGAATTTGTAACAGAAAGATCTCAAGCCCTAGAAAAATCTGTCAGTAACCTGGAG 1500
Db LeuGlnAlaAlaMetHisAspAspLeuLysGluValGluLysSerIleThrAsnLeuGlu 528
Qy GAATCCCTAACCTCTTATCTGAAGTAGTCTACAGAAATAGAGAGGGTTAGATTATTA 1560
Db LysSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 548
Qy TTTCTAAAGAAAGAGGATATGTAGTCCTTGAAGGAGGAATGCTGTTTTTATGTGGAT 1620
Db PheLeuLysGluGlyGlyLeuCysAlaAlaLeuLysGluGlyCysPheTyrAlaAsp 568
Qy CATTACGGGGCCATCAGAGACTCCATGAACAAAGCTTAGAGAAAGGTTGGAGAACGCTCGA 1680
Db HisThrGlyLeuValArgAspSerMetAlaLysLeuArgPheArgLeuSerGlnArgGln 588
Qy AGGAAAAGGAACTACTCAAGGTGGTTTGAGGGATGGTTCAACAGGTCTCTTTGGTTG 1740
Db LysLeuPheGluSerGlnGlnGlyTrpPheGluGlyLeuPheAsnLysSerProTrpPhe 608
Qy GCTACCTCTCTCTGCTTTAAACAGGACCCTTAATAGTCTCTCTCTCTTACTCACAGTT 1800
Db ThrThrLeuLeuSerThrIleMetGlyProLeuIleLeuLeuLeuIleLeuLeuPhe 628
Qy GGGCCATGTATTATTAACAAGTTAATTCCTTTCATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db GlyProCysIleLeuAsnArgLeuValGlnPheIleLysAspArgIleSerValValGln 648
Qy ATCATGTACTTAGACACAGTAC 1884
Db AlaLeuValLeuThrGlnGlnTyr 656

RESULT 9
B43491
env polyprotein - T3651/B murine leukemia virus
C:Species: T3651/B murine leukemia virus
C:Date: 12-Jan-1993 #sequence_revision 12-Jan-1993 #text_change 09-Jul-2004
C:Accession: B43491
R:Baylac-Kalabokias, H.; Astier-Gin, T.; Moynet, D.; Hernould, M.; Mamoun, R.; Legrand, Virus Res. 18, 117-134, 1991
A:Title: A new leukemogenic retrovirus isolated from tumor cells derived from a radio-in A:Reference number: A43491; MUID:91253260; PMID:2042396
A:Accession: B43491
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-689 <BAY>
A:Cross-references: UNIPROT:Q9Q9A5; UNIPARC:UPI00000178638; GB:X59001

C:Superfamily: type C retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protein

Alignment Scores:
Pred. No.: 1-82e-98 Length: 689
Score: 1335.00 Matches: 296
Percent Similarity: 57.7% Conservatives: 92
Best Local Similarity: 44.0% Mismatches: 181
Query Match: 38.5% Indels: 104
DB: 2 Gaps: 22

US-10-723-552-3_COPY_5620_7533 (1-1914) x B43491 (1-689)

Qy 97 CTTACTCTGTCAATAAACCTCTCAGACTAATGGTATGCGCATAGGAGACAGCCTGAACCTCC 156
Db 39 IleValLeuLeuIleLeuGlyGlyValAsnProValAlaLeuGlyAsnSer-----Pro 56
Qy 157 CATAAACCCCTTATCTCTCACCCTGTTAATTAATCACTCCGCCACAGGATTAATATCAAC 216
Db 57 HisGlnValPheAsnLeuSerTrpGluValThrAsnGlyGlyLeuGlu---ThrValTrp 75
Qy 217 AACACTCAAGGGAGGCTCTTTTAGGAACCTGGTGGCTGATCTATATAC-----GTTTGC 270
Db 76 AlaIleThrGlyAsnHisProLeuTrpThrTrpProAspLeuThrProAspLeuCys 95
Qy 271 CTCAGATCAGTTATT---CCTAGTCTG-----ACCTCACCCC 303
Db 96 MetLeuAlaLeuHisGlyProSerTyrTrpGlyLeuGluTyrArgAlaProPheSerPro 115
Qy 304 CCA-----306
Db 116 ProGlyProProCysCysSerGlySerAsnAspSerThrSerGlyCysSerArgAsp 135
Qy 306-----306
Db 136 CysGluGluProLeuThrSerTyrThrProArgCysAsnThrAlaTrpAsnArgLeuLys 155
Qy 307 -----GATATCCTCCATGCTCAC-----GGATTTTATGTTTCCCGAGGACCA----- 348
Db 156 LeuSerLysValThrHisAlaHisAsnGluGlyPheTyrValCysProGlyProHisArg 175
Qy 349 CCAATAATCGAAAAACATTCGGGAAATCCACAGAGATTTCTTTTGAACAAATGGAACCTGT 408
Db 176 ProArgTrpAlaArgSerCysGlyGlyProGluSerPheTyrCysAlaSerTrpGlyCys 195
Qy 409 GTAACCTCTAATGATGATATCGAATGCGAATGCGCAACCTCTCAGCAGGATAGGTAAGTTT 468
Db 196 GluThrThrGlyArgAlaSerTrpLys---ProSerSerSerTrpAspTyrIleThrVal 214
Qy 469 TCTTATGTCAACACCTATACAGCTCTGGACAATTTAATTACCTGACCTGGATTAGAACT 528
Db 215 Ser-----AsnAsnLeuThrAlaAsp-----Gln 222
Qy 529 GGAAGCCCCAAGTGTCTCTCCTTCAGAC---CTAGATTACCTAAAAATAAGTTTCACTGAG 585
Db 223 AlaThrProAlaCysLysGlyAsnLysTrpCysAsnSerLeuThrIleArgPheThrSer 242
Qy 586 AAGGAAAAACAAGAAATATCCTAAAATGGTAAATGGTATGTCCTTGGGGAATGGTATAT 645
Db 243 PheGlyLysGlnAlaThr-----SerTrpValThrGlyHisTrpTrpGlyLeuArgLeu 260
Qy 646 TATGGAGGCTCGGGTAAACAAACACAGGCTCCATTTCAACTATTTCGCTCAAAAATAAACAG 705
Db 261 Tyr---ValSerGlyHisAspProGlyLeuIlePheGlyIleArgLeuLysIleThrAsp 279
Qy 706 CTGAGCCTCAATGGCTATPAGACCAATACCGGTCTTTCAGCGGTCAAGACCC----- 759
Db 280 LeuGlyProArgValProIleGlyProAsnProValLeuSerAspArgArgProProSer 299
Qy 760 -----CCACCCCAAGGACGAGCCATCTCTTAACATAACTTCTGATCAGACCCC 810
Db 300 ArgProArgProThrArgSerPro---ProSerSerAsn-----SerThrPro 314

Db 223 ThrProAlaCysGlyAsnLysTrpCysAsnSerLeuThrIleArgPheThrSerPhe 242
QY 589 GGAACCAAGAAATATCTAAATCGGTAAATGGTATGCTCTGGGGAATGTATATAT 648
Db 243 GlyLysGlnAlaThr-----SerTrpValThrGlyHisTrpTrpGlyLeuArgLeuTyr 260
QY 649 GGAGCTCGGTAAACACACAGCTCCATCTTAACCTATTCGCTCAAAATAACACAGCTG 708
Db 261 ---ValSerGlyHisAspProGlyLeuIlePheGlyIleArgLeuLysIleThrAspLeu 279
QY 709 GAGCTCCCAATGGCTATAGGACCAATACGGTCTTACGGGTCAAAGACCC----- 759
Db 280 GlyProArgValProIleGlyProAsnProValLeuSerAspArgArgProSerArg 299
QY 760 -----CCAACCCAGGACACGACATCTCTAATACTTCGGATCAGACCCCACT 813
Db 300 ProArgProThrArgSerProProSerAsn-----SerThrProThr 314
QY 814 GAGCTACACAGCAG-----ACTAAATGGGGGCAAACTTTTACG 855
Db 315 GluThrProLeuThrLeuProGluProProAlaGlyValGluAsnArgLeuLeuAsn 334
QY 856 CTCATCCAGGAGCTTTTCAAGCTCTTAAGCTCCAGACTCCAGAGGCTACCTCTTCTGT 915
Db 335 LeuValLysGlyAlaTyrGlnAlaLeuAsnLeuThrSerProAspLysThrGlnGluCys 354
QY 916 TGGCTATCTTAGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAGAGGGAATTC 975
Db 355 TrpLeuCysLeuValSerGlyProProTyrTrpGluGlyValAlaValLeuGlyThrTyr 374
QY 976 AATGTGACAAAGAACAATAGACACCATGCATCGGGATCCCAAAATAAGCTTACCTT 1035
Db 375 SerAsnHisThrSerAlaProAlaAsnCysSerValAlaSerGlnHisLysLeuThrLeu 394
QY 1036 ACTGAGGTCTTCGAAAAGGCACCTGCATAGCAAGGTTCCCATCCACCAACACCTT 1095
Db 395 SerGluValThrGlyGlnGlyLeuCysIleGlyAlaValProLysThrLeuGlnValLeu 414
QY 1096 TGTAAACCACTGAAGCCCTTTAATCAAACTCTGAGAGTCAATATCTGCTACCTGGTTAT 1155
Db 415 CysAsnThrThrGln-----LysThrSerAlaGlySerTyrTyrLeuAlaAlaProThr 432
QY 1156 GACAGGTGGTGGCATGTAATCTGATTAACCCCTGTGTTCACCTCGTGGTTTAAAC 1215
Db 433 GlyThrIleTrpAlaCysSerThrGlyLeuThrProCysIleSerThrIleLeuAsn 452
QY 1216 CAATCAAGATTTTTCATTTATGTCCTCAAAATGTTCCCGAGTGTATTACTATCCCGAA 1275
Db 453 LeuThrThrAspTyrCysValLeuValGluLeuTrpProArgValThrTyrHisSerPro 472
QY 1276 AAAGCAATCTTGATGAATATGACTACAGAAATCATCGACAAAGAGAGAACCCATATCT 1335
Db 473 SerTyrValTyrHisGlnPheGlu---ArgArgAlaLysTyrLysArgGluProValSer 491
QY 1336 CTGACACTTGTGTGATGCTCGGA-----CTTGAGTGGCAGCAGGTGTAGGA 1383
Db 492 LeuThrLeuAlaLeuLeuLeuGlyGlyLeuThrMetGlyGlyIleAlaAlaGlyValGly 511
QY 1384 ACAGGACAGCTGCCCTGGTCAACGGACCAAGCAGCTAGAAACAGGACTAGTAACCTA 1443
Db 512 ThrGlyThrThrAlaLeuValAlaThrGlnGln-----PheGlnGlnLeu 526
QY 1444 CATCGAATTTGAACAGAGATCTCCAAGCCCTAGAAAATCTGTCAAGTAACCTGAGGAA 1503
Db 527 GlnAlaAlaMetHisAspAspLeuLysGluValGluLysSerIleThrAsnLeuGluLys 546
QY 1504 TCCCTAACCTCTTATCTGAAGTATGCTTACAGAAATGAAGAGGTTAGATTATTATTT 1563
Db 547 SerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuPhe 566
QY 1564 CTAAGACAGGAGATTTATGTTAGCTTGNAGGAGGAATGCTGTTTTTATGTCGATCAT 1623
Db 567 LeuLysGluGlyGlyLeuCysAlaAlaLeuLysGluCysPheTyrAlaAspHis 586

QY 1624 TCAGGGCCATCAGAGACTCCATGAACAAGCTTACAGAAAGCTTGAGAAAGCGTCAAGG 1683
Db 587 ThrGlyLeuValArgAspSerMetAlaLysLeuArgGluArgLeuSerGlnArgGlnLys 606
QY 1684 GAAAGGAAACTACTCAAGGGTGGTTGAGGATGGTTCAACAGGTCTCTTTGGTTGGCT 1743
Db 607 LeuPheGluSerGlnGlnGlyTrpPheGluGlyLeuPheAsnLysSerProTrpPheThr 626
QY 1744 ACCCTACTTCTGCTTTAAACAGGACCTTAATAGTCCTCTCTCTACTACACTGTTGG 1803
Db 627 ThrLeuIleSerThrIleMetGlyProLeuIleIleLeuLeuLeuLeuPheGly 646
QY 1804 CCATGTATTATTAAACAGTTAAATTCCTTTCATTAGAGAACGAATAAGTCAGTCAGATC 1863
Db 647 ProCysIleLeuAsnArgLeuValGlnPheIleLysAspArgIleSerValValGlnAla 666
QY 1864 ATGCTACTTACACACAGTACCAAGCCCGCTAGC 1899
Db 667 LeuValLeuThrGlnGlnTyrHisGlnLeuLysSer 678
RESULT 11
VCVWEM
env polyprotein - Moloney murine leukemia virus
N:Contains: knob protein gp70; R protein; spike protein p15e
C:Species: Moloney murine leukemia virus
C>Date: 01-Sep-1981 #sequence revision 24-Sep-1981 #text_change 09-Jul-2004
C:Accession: A93265; A93235; A93848; A03983
R:Shinnick, T.M.; Lerner, R.A.; Sutcliffe, J.G.
Nature 293, 543-548, 1981
A:Title: Nucleotide sequence of Moloney murine leukaemia virus.
A:Reference number: A93265; MUID:82035843; PMID:6169994
A:Accession: A93265
A:Molecule type: genomic RNA
A:Residues: 1-565 <SH>
A:Cross-references: UNIPROT:P03385; UNIPARC:UPI00000134DC; GB:J02255; GB:J02256; GB:J0.
A:Experimental source: clone pMLV-1
R:Sutcliffe, J.G.; Shinnick, T.M.; Green, N.; Liu, P.T.; Niman, H.L.; Lerner, R.A.
Nature 287, 801-805, 1980
A:Title: Chemical synthesis of a polypeptide predicted from nucleotide sequence allows
A:Reference number: A93235; MUID:81052384; PMID:6159543
A:Accession: A93235
A:Molecule type: genomic RNA
A:Residues: 496-665 <SUT1>
A:Cross-references: UNIPARC:UPI0000174A47
A:Experimental source: provirus
R:Sutcliffe, J.G.; Shinnick, T.M.; Verma, I.M.; Lerner, R.A.
Proc. Natl. Acad. Sci. U.S.A. 77, 3302-3306, 1980
A:Title: Nucleotide sequence of Moloney leukemia virus: 3' end reveals details of repli
A:Reference number: A93848; MUID:81013872; PMID:6251454
A:Accession: A93848
A:Molecule type: DNA
A:Residues: 484-662, 'CEF' <SUT2>
A:Cross-references: UNIPARC:UPI0000174A48
A:Experimental source: provirus, clone pMLV-201
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane prote
F:34-469/Product: knob protein gp70 #status predicted <KNB>
F:470-649/Product: spike protein p15e #status predicted <SPK>
F:650-665/Product: R protein #status predicted <RPT>
F:45,199,356,358,398,434/Binding site: carbohydrate (Asn) (covalent) #status predicted
Alignment Scores:
Pred. No.: 3,43e-97 Length: 665
Score: 1319.00 Matches: 291
Percent Similarity: 55.7% Conservative: 86
Best Local Similarity: 43.0% Mismatches: 198
Query Match: 38.0% Indels: 102
DB: 1 Gaps: 20
US-10-723-552-3_COPY_5620_7533 (1-1914) x VCVWEM (1-665)

Qy	49	CCGAAAGACTGAAATCCCTTTAAAGCTTCGCCTCCATCGCGCTGGTTCCTTACTCTGTCA	108
Db	16	ProArgGlyProLeuLeuIleProLeuLeuMetLeuArgGlyValSerThrAlaSer	35
Qy	109	ATAACCTCTCAGACTAAATGGTATGCGCATAGAGACAGCGCTGAATCCCATTAACCCCTTA	168
Db	36	ProGlySer-----SerProHisGlnValTyr	44
Qy	169	TCTCTCACTGGTTAAATTAAGTACTCCGCGACAGGTATTAATATCAACAACACTCAAGG	228
Db	45	AsnIleThrTyrGluValThr---AsnGlyAspArgGluThrValTyrProAlaThrSerGly	63
Qy	229	GAGGCTCCTTTAGGAACCTGGTGGCTGATCATAC-----GTTTGGCTC-----	273
Db	64	AsnHisProLeuTyrThrTyrProAspLeuThrProAspLeuCysMetLeuAlaHis	83
Qy	274	-----AGATCAGTTATTCCTAGTCTGACCTCACCC	303
Db	84	HisGlyProSerTyrTyrGlyLeuGluTyrGlnSerProPheSerSerProGlyPro	103
Qy	304	CCA-----GATATCTCTCCAT	318
Db	104	ProCysCysSerGlyCysSerProGlyCysSerArgAspCysGluGluProLeuThr	123
Qy	307	-----GATATCTCTCCAT	318
Db	124	SerLeuThrProArgCysAsnThrAlaTyrAsnArgLeuLysLeuAspGlnThrThrHis	143
Qy	319	GCT-----CAGCGATTTTATGTTTCCCGACAGCA-----CCAATTAATCGAAACAT	366
Db	144	LysSerAsnGluGlyPheTyrValCysProGlyProHisArgProArgGluSerLysSer	163
Qy	367	TGCGAAATCCAGAGATTTCTTTGTAAACAATGGAACGTGTGAACCTCTAATGATGA	426
Db	164	CysGlyGlyProAspSerPheTyrCysAlaTyrTyrGlyCysGluThrThrGlyArgAla	183
Qy	427	TATTGGAAATGCCAACCTCTCAGCAGGATAGGGTAAGTTTTTCTTATGTCAACACCTAT	486
Db	184	TyrTrpLys---ProSerSerSerTrpAsp-----PheIleThrValAsnAsnAsn	199
Qy	487	ACCAGCTCTGGACAATTTAATTACCTGCCTGGATTAGAACTGGAAGCCCAAGTGCTCT	546
Db	200	LeuThrSerAspGln-----AlaValGlnValCysLysAspAsnLysTrpCysAsn	216
Qy	547	CCTTCAGACCTAGATTACCTAAAAATAAGTTTCACTGAGAAAGAAACAAGAAATAATC	606
Db	217	Pro-----LeuValIleArgPheThrAspAlaGlyArgArg-----Val	229
Qy	607	CTAAATGGTAAATGGTATGTCCTTGGGAATGGTATATTATGAGAGGCTCGGGTAAACAA	666
Db	230	ThrSerTyrThrThrGlyHisTyrTrpGlyLeuArgLeuTyr---ValSerGlyGlnAsp	248
Qy	667	CCAGGCTCCATTCTTAACATTTCGCTCAAAATAAACCCAGCTCGAGCTCCAATGGCTATA	726
Db	249	ProGlyLeuThrPheGlyIleArgLeuArgTyrGlnAsnLeuGlyProArgValProIle	268
Qy	727	GGACCAAAATACGGTCTTGACGGGTCAAGACCC-----CCAACCAAGGACCAAGA	777
Db	269	GlyProAsnProValLeuAlaAspGlnGlnProLeuSerLysProLysProValLysSer	288
Qy	778	CAATCCTCTAACAATACTCTCGATCAGACCCCACTGAGTCTAACAGCAGGACTAAATG	837
Db	289	ProSerValThrLysProProSerGlyThrProLeuSerProThrGlnLeuProProAla	308
Qy	838	GGGGCA-----AAACTTTTAAAGCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCAG	891
Db	309	GlyThrGluAsnArgLeuAsnLeuValAspGlyAlaTyrGlnAlaLeuAsnLeuThr	328
Qy	892	ACTCCAGAGGCTACTCTTCTTTGTGGCTATGCTTAGCTTCGGGGCCCACTTACTATGAA	951
Db	329	SerProAspLysThrGlnGluCysTrpLeuCysLeuValAlaGlyProProTyrTyrGlu	348

QY	952	GGAAATGCTAGAGAGGGGAAATTCATAGTGCACAAAAGACATAGAGACCCATCCATCG	1011
DB	349	GlyValAlaValLeuGlyThrTySerAsnHisThrSerAlaProAlaAsnCySerVal	368
QY	1012	GGATCCCAAAATAAGCTTACCTTACTTGAGGTTCCTCGAAAGAGCCACCTGCATAGGAAG	1071
DB	369	AlaserGlnHisIysLeuThrLeuSerGluValThrGlyGlnGlyLeuCySilleGlyAla	388
QY	1072	GTTCCCCCATCCCAACCAACCTTGTGTAAACACACTGAAGCCCTTAAATCAAACTCTGAG	1131
DB	389	ValProIysThrHisGlnAlaLeuCyAsnThrThr-----GlnThrSerSer	404
QY	1132	-----AGTCAATATCTGGTACCTGGTTATGACAGGTGGTGGGCATGTAATACTGATTA	1185
DB	405	ArgGlySerTyTyLeuValAlaProThrGlyThrMetTrpAlaCysSerThrGlyLeu	424
QY	1186	ACCCCTGTGTTCCACCTTGGTGTTTTAAACAAACTAAAGATTTTGTGCATTATGGTCCAA	1245
DB	425	ThrProCysIleSerThrThrIleLeuAsnLeuThrThrAspTyTyCysValLeuValGlu	444
QY	1246	ATTCTTCCCGAGTGTTACTATCCCGAAAAAGCAATCCTTGATGAATATGACTACAGA	1305
DB	445	LeuTrpProArgValThrTyHisSerProSerTyTyValTyGlyLeuPheGlu---Arg	463
QY	1306	AATCATCGACAAAGAGAGAACCCATATCTCTGACACTGTGCTGTGATGCTCGGA	1359
DB	464	SerAsnArgHisIysArgGluProValSerLeuLeuAlaLeuLeuLeuGlyGlyLeu	483
QY	1360	-----CTTCGAGTGGCAGCAGGTGTAGGAAACAGGACAGCTCCCTCGTGCACGGGACCA	1413
DB	484	ThrMetGlyGlyIleAlaGlyIleGlyThrGlyThrThrAlaLeuMetAlaThrGln	503
QY	1414	CAGCAGCTAGAAAACAGGACTTAGTAACCTACATCGAATTCTAAACAGAAAGATCTCCAGCC	1473
DB	504	Gln-----PheGlnGlnLeuGlnAlaAlaValGlnAspAspLeuArgGlu	518
QY	1474	CTAGAAAAATCTGTCACTAACCTGGAGGAATCCTTAACCTCTTATCTGAAGTAGTCCTTA	1533
DB	519	ValGluIysSerIleSerAsnLeuGluLysSerLeuThrSerLeuSerGluValValLeu	538
QY	1534	CAGAAATAGAGGGTTAGATTATTTATTTCTAAAGAGGAGGATTATGTGTAGCCTTG	1593
DB	539	GlnAsnArgArgGlyLeuAspLeuLeuPheLeuIysGlyGlyLeuCySalleAlaLeu	558
QY	1594	AAGGAGAAATCTGTTTTTATGTGGATCATTTACGGGGCCATCAGAGACTCCATGAACAAG	1653
DB	559	LysGluGluCysCysPheTyAlaAspHisThrGlyLeuValArgAspSerMetAlaLys	578
QY	1654	CTTAGAAGAAAGGTTGGAGAACGTCGACGGGAAAGGAAACTACTCAAGGGTGGTTGAG	1713
DB	579	LeuArgGluArgLeuAsnGlnArgGlnLysLeuPheGluSerThrGlnGlyTyTrpPheGlu	598
QY	1714	GGATGGTTCAACAGGTCTCTTTGGTGGCTTACCTACTTTCTGCTTTTAAACAGACCCCTTA	1773
DB	599	GlyLeuPheAsnArgSerProTyPheThrThrLeuIleSerThrIleMetGlyProLeu	618
QY	1774	ATAGTCTCTCTCTGTGTACTCACAGTTGGGCCCATGTATTATTAAACAGTTTAATTCGCTTC	1833
DB	619	IleValLeuLeuMetIleLeuLeuPheGlyProCysIleLeuAsnArgLeuValGlnPhe	638
QY	1834	ATTAGAGAACGAATAAGTCGAGTCACAGATCATCGTACTTAGACAAACAGTAC	1884
DB	639	ValIysAspArgIleSerValValGlnAlaLeuValLeuThrGlnGlnTyTr	655

RESULT 12

RESCEI 12
VCMVHL

VCMVHL
env po]vprotein precursor - HoMuLV murine leukemia virus

N:Contains: knob protein gp76: spike protein p15E

N; contains: knob protein gp6; spike pr
C; species: HcMuLV murine leukemia virus

C;Species: HoMuLV murine leukemia virus
A;Note: host Mus hortulanus (European mouse)

A;Note: host *Mus hortulanus* (European mouse)
C.Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text change 09-Jul-2004

C;Date: 31-Dec-1990
C:Accession: B32594

C;Accession: B32594
R:Vovtek P.: Kozak C.A.

Virology 173, 58-67, 1989
 A:Title: Nucleotide sequence and mode of transmission of the wild mouse ecotropic virus,
 A:Reference number: A32594; MUID:90051094; PMID:2554579
 A:Accession: B32594
 A:Molecule type: DNA
 A:Residues: 1-666 <VOY>
 A:Cross-references: UNIPROT:P21436; UNIPARC:UPI000012A03F
 A:Note: the authors translated the codon CTT for residue 451 as Pro
 C:Genetics:
 A:Gene: env
 C:Superfamily: type C retrovirus env polypeptide
 C:Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protein
 F:1-30/Domain: signal sequence #status predicted <SIG>
 F:31-466/Product: knob protein gp76 #status predicted <KPG>
 F:467-666/Product: spike protein p15E #status predicted <SPP>
 F:473-489/Domain: transmembrane #status predicted <TM1>
 F:607-623/Domain: transmembrane #status predicted <TM2>
 F:42,197,290,324,356,363,431,599/Binding site: carbohydrate (Asn) (covalent) #status pre

Alignment Scores:

Pred. No.:	1.03e-96	Length:	666
Score:	1313.00	Matches:	300
Percent Similarity:	57.4%	Conservative:	95
Best Local Similarity:	43.6%	Mismatches:	182
Query Match:	37.9%	Indels:	111
DB:	1	Gaps:	23

us-10-723-552-3_copy_5620_7533 (1-1914) x VCMVHL (1-666)

Qy	97	CTTACTCTGTCATTAACCTCTCAGACTAATGGT-----	129
Db	6	LeuProlySerIleLeuAspLeuThrAsnProTrpGlyProIleIleLeuGlyIleLeu	25
Qy	130	ATGCGCATAGGAGCAGCTG-----AACTCCCATAAACCTTATCTCAGCTGG	180
Db	26	IleMetLeuGlyAlaLeuGlySerProHisLeuValPheAsnLeuThrTrp	45
Qy	181	TTAATTACTGCTCGGCACAGGTATTAAATACACACACTCAAGGGAGCTCTTTA	240
Db	46	GluValTyrAsnGlnGluTyrGlu---ThrValTrpAlaThrSerGlySerHisProLeu	64
Qy	241	GGAACTGGTGGCTGATCTATAC-----GTTGGCTCAGATCAGTATT--	285
Db	65	TrpThrTrpTrpProThrLeuProAspLeuCysMetLeuAlaGlnLeuAlaLysPro	84
Qy	286	-----CCTAGTCTGACCTCACCCTCAGAT-----	309
Db	85	SerTrpGlyLeuSerAspTyrProProTyrSerLysProGlyProProCysCysThr	104
Qy	309	-----	309
Db	105	ThrAspAsnAsnProProGlyCysSerArgAspCysAsnGlyProLeuThrTyrLeuThr	124
Qy	310	-----ATCTCCATGCTCAC-----	324
Db	125	ProArgCysSerThrAlaTrpAsnArgLeuLysLeuValLeuThrThrHisLeuAsn	144
Qy	325	---GGATTATGTTTCCAGGACCA-----CCAAATAATGAAACATTCGGGAAT	375
Db	145	GlnGlyPheTyrValCysProGlyProHisArgProArgHisAlaArgAsnCysGlyGly	164
Qy	376	CCAGAGATTCTTTTGTAAACATGGAACCTGTGTAACTCTTAATGATGATATTGGAA	435
Db	165	ProAspAspPheTyrCysAlaHisTrpGlyCysGluThrThrGlyGlnAlaTyrTrpLys	184
Qy	436	TGGCCAACTCTCAGCAGGATAGGTAAATTTTCTATGTCACACCTATACAGCTCT	495
Db	185	---ProSerSerTrpAspTyrIleArgValSer-----AsnAsnAlaSerSerSer	201
Qy	496	GGACAATTTAATTACCTGACTGATAGACTGGAAGCCCAAGCTCTCTCTCAGAC	555
Db	202	Asp-----AlaThrAlaCysLysAsnAsnAsnTrpCysSerPro-----	215

Qy	556	CTAGATTACCTAAATAAAGTTTTCACCTGAGAAAGAAACCAAGAAATATCTCTAAATGG	615
Db	216	-----LeuAlaIleSerPheThrAspProGlyLysArgAlaThr-----SerTrp	230
Qy	616	GTAATGGTATGCTTGGGGAATG---GTATATTATGAGGCTCGGGTAAACACAGGC	672
Db	231	ThrSerGlyPheThrTrpGlyLeuArgLeuTyrIleSerGly-----HisProGly	247
Qy	673	TCCATTCTTAACATTTCGCTCAAATAAACACAGCTGGAGCTCCCAATGGCTATAGACCA	732
Db	248	LeullePheGlyValArgLeuLysIleSerAspLeuGlyProArgValProIleGlyPro	267
Qy	733	AATACGGCTTTCAGCGGTCAAAGACCCCAACCAAGGACCA-----GGACCA	780
Db	268	AsnProValLeuSerGluGlnArgProProSerGlnProGluProAlaArgLeuProPro	287
Qy	781	TCCTCTAACATAACTTCTGGATCAGACCCCACTGAGTCTAACAGCAGCACTAAATGGG	840
Db	288	SerSerAsnLeuThrGlnGlyGlyThrProSerAlaProThrGlyProProGlnGly	307
Qy	841	GCA-----AAACTTTTTCAGCTCATCCAGGAGCTTTCAAGCTCTTAACTCCACGACT	894
Db	308	ThrGlyAspArgLeuLeuAspLeuValGlnGlyAlaTyrGlnAlaLeuAsnAlaThrSer	327
Qy	895	CCAGAGCTACCTCTCTTGTGGCTATGCTTTCGGGCCCACTTACTATGAAGGA	954
Db	328	ProAspLysThrGlnGlyCysTrpLeuCysLeuValSerSerProProTyrTyrGlyGly	347
Qy	955	ATGCTAGAGAGGGAATTCATGTGACAAAGAACATGAGACCAATGCACATGGGGA	1014
Db	348	ValAlaValValGlyProTyrSerAsnHisThrThrAlaProAlaAsnCysSerAlaAsp	367
Qy	1015	TCCCAAATAAGCTTACCTTACTGAGTTCCTGGAAGGACCTGCATAGGAAGGTT	1074
Db	368	SerGlnHisLysLeuThrLeuSerGluValThrGlyLys---ProLeuProArgLysGly	387
Qy	1075	CCCCATCCCAACCAACCTTTGTAAACCACTGAAGCTTTAATCAACCTCTGAGAGT	1134
Db	387	erGlnAspProProGlyProValGlnTyrHis-----SerGly-AlaArgGln	402
Qy	1135	CAATATCTGTTACCTGTTATGACAGG-----TGGTGGGATGTAATATCTGGATTACC	1189
Db	403	LysTyrSerLeuSerGlyGlySerArgGlyThrMetTrpAlaCysAsnThrGlyLeuThr	422
Qy	1189	CCTTGTCTTCCACTTGGTTTTTAAACCAATAAGATTTTTCATTATGCTCCAAAT	1248
Db	423	ProCysLeuSerThrAlaValLeuAsnLeuThrThrAspTyrCysValLeuValGluLeu	442
Qy	1249	GTTCCCGAGTGTATTACTATATCCGAAAAAGCAATCTTGTATGAATATGACTACAGAAAT	1308
Db	443	TrpProArgValThrTyrHisSerLeuAspPheValTyrArgGlnValGlyArgThr	462
Qy	1309	CATCGACAAAGAGAGAACCATATCTCTGACACTTCTGCTGTGATGCTCGGA-----	1359
Db	463	---ArgTyrGlnArgGluProValSerLeuThrLeuAlaLeuLeuLeuGlyGlyLeuThr	481
Qy	1360	---CTTCGAGTGCAGCAGGTGTAGGAACAGGACAGCTCCCTGGTCCAGGACCAAG	1416
Db	482	MetGlyGlyIleAlaAlaGlyValGlyThrGlyThrSerAlaLeuValLys---ThrGln	500
Qy	1417	CAGCTAGAAACAGGACTTAGTAACTACATCGAATTTGTAAACAGAAAGATCTCCAGCCCTA	1476
Db	501	GlnPheGlu-----GlnLeuHisAlaAlaIleGlnAlaAspLeuLysGluVal	516
Qy	1477	GAAAAATCTCTCAGTAACCTGGAGGAATCCCTAACCTCTTATCTGAAGTAGTCTCTACAG	1536
Db	517	GluSerSerIleThrAsnLeuLysSerLeuThrSerLeuSerGluValValLeuGln	536
Qy	1537	AATAGAGAGGGTTAGATTATTATTCTAAAGAGGAGGATTATGCTAGCTCTCAAG	1596
Db	537	AsnArgArgGlyLeuAspLeuLeuGlyGlyLeuCysAlaLeuLys	556
Qy	1597	GAGGAATGCTGTTTTTTTATGTGGATCATTCAGGGGCCATCAGAGACTCCATGAACAAGCTT	1656

```
Db 557 GluGluCysCysPheTyrAlaAspHisThrGlyLeuValArgAspSerMetAlaLysLeu 576
|||||
Qy 1657 AGAGAAAGCTGGAGAACGCTGAGAGGAAAGAAAGAACTACTCAAGGCTGGTTGAGGGA 1716
|||||
Db 577 ArgGluArgLeuAsnGlnArgGlnLysLeuPheGluAlaGlyGlnGlyTrpPheGluGly 596
|||||
Qy 1717 TGGTTCAACAGCTCTCTTTGGTGGCTACCTACTCTTCTGCTTAAACAGGACCTTAATA 1776
|||||
Db 597 LeuPheAsnArgSerProTrpPheuThrThrLeuIleSerThrIleMetGlyProLeuIle 616
|||||
Qy 1777 GTCCTCTCTCTGTTACTCACAGTTGGGCATGTATTATTAAACAAGTTAAATGCTTCATT 1836
|||||
Db 617 IleLeuLeuLeuIleLeuMetPheGlyProCysIleLeuAsnArgLeuValGlnPheVal 636
|||||
Qy 1837 AGAGAACGAATAAGTGCAGTCAGATCATGCTAGTACTTAGACACAGTACCAAGCCGCTCT 1896
|||||
Db 637 LysAspArgIleSerValGlnAlaLeuValLeuThrGlnGlnTyHieGlnLeuLys 656
|||||
Qy 1897 AGCAGGGAAGCTGGCCGC 1914
|||||
Db 657 ProLeuGluHisGlyArg 662
|||||

RESULT 13
VCMVRV
env polyprotein precursor - Rauscher mink cell focus-forming virus
N:Contains: coat protein p12E; coat protein p15E; knob protein gp70
C:Species: Rauscher mink cell focus-forming virus
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A03990
R:Voigt, M.; Haggbloom, C.; Swift, S.; Haas, M.
J. Virol. 55, 184-192, 1985
A:Title: Envelope gene and long terminal repeat determine the different biological prop
A:Reference number: A93011; MUID:85237696; PMID:4009793
A:Accession: A03990
A:Molecule type: DNA
A:Residues: 1-640 <VOG>
A:Cross-references: UNIPROT:P06445; UNIPARC:UPI000012A047; GB:M10100; NID:9332068; PIDN:
C:Gene: env
C:Superfamily: type C retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protei
F:1-32/Domain: signal sequence #70 status predicted <SIG>
F:33-443/Product: knob protein gp70 #status predicted <KBP>
F:444-640/Product: coat protein p15E #status predicted <PFE>
F:444-623/Product: coat protein p12E #status predicted <PTE>
F:43,58,300,332,339,372,408,576/Binding site: carbohydrate (Asn) (covalent) #status pred

Alignment Scores:
Pred. No.: 7,06e-95 Length: 640
Score: 1290.00 Matches: 277
Percent Similarity: 59.9% Conservative: 104
Best Local Similarity: 43.6% Mismatches: 171
Query Match: 37.2% Indels: 84
DB: 1 Gaps: 19

US-10-723-552-3_COPY_5620_7533 (1-1914) x VCMVRV (1-640)
Qy 130 ATGCGCATAGGACAGACGCTG-----AACTCCCAATAAACCTTATCTCTCACTGG 180
|||||
Db 27 IleArgAlaGlyValSerValGlnHisAspSerProHisLysValPheAsnValThrTrp 46
|||||
Qy 181 TTAATTACTGACTCCGGCAGGATTAATATACACACACTCAAGGAGGAGCTCTCTTA 240
|||||
Db 47 ArgValThrAsnLeuMetThrGlyGlnThrAlaAsnAlaThr-----SerLeuLeu 63
|||||
Qy 241 GGAACCTCG-----TGGCTCATCTATACGTTTGGCTCAGATCAGTTATTCCTAGT 291
|||||
Db 64 GlyThrMetThrAspAlaPheProLysLeuTyrPheAspLeuCysAspLeuValGlyAsp 83
|||||
Qy- 292 CTG-----ACCTCACCCCGCATATC----- 312
|||||
Db 84 TyrTrpAspAspProGluProAspIleGlyAspGlyCysArgThrProGlyGlyArgArg 103
|||||
```

```
Qy 313 ---CTCCATGCTCAGGATTTTATGTTTCCAGGACCCACCAATAATATGGAACAACAT--- 366
|||||
Db 104 ArgThrArgLeuTyrAspPheTyrValCysProGly-----HisThr 117
|||||
Qy 367 -----TCGGAAATCCAGAGATTTCTTTTGTAAACAATGGAACCTGTGTAACC 414
|||||
Db 118 ValProIleGlyCysGlyGlyProGlyGlyTyrCysGlyLysTrpGlyCysGluThr 137
|||||
Qy 415 TCTAATGATGATATTGGAATGCGCAACCTCTCAGCAGGATAGGTAAGTTTCTTAT 474
|||||
Db 138 ThrGlyGlnAlaTyrTrpLys---ProSerSerSerTrpAspLeuIleSerLeuLysArg 156
|||||
Qy 475 GTCAACACC-----TATACCAGCTCTGGACAAATTAATTAACCTG 513
|||||
Db 157 GlyAsnThrProLysAspGlnGlyProCysTyrAspSerSerValSerSerAsp----- 174
|||||
Qy 514 ACCTGGATTAGAACTGGAAGCCCC-----AAGTGTCTCTCTTCAGACCTAGATTACCTA 567
|||||
Db 175 -----IleLysGlyAlaThrProGlyGlyArgCysAsnPro-----Leu 187
|||||
Qy 568 AAAATAAGTTTCACTGAGAAAGAAACAAAGAAATATCTTAAATGGTAAATGGTATG 627
|||||
Db 188 ValLeuGluPheThrAspAlaGlyLysLysAlaSer-----TrpAspGlyProLys 204
|||||
Qy 628 TCTTGGGATGGTATATTATGAGGCTCGGGTAAACAACACCGAGCTCCATTTCTAACTATT 687
|||||
Db 205 ValTrpGlyLeuArgLeuTyrArgSerThrGlyThrAspProValThrArgPheSerLeu 224
|||||
Qy 688 CGCCTCAAAATAAACACAGCTGGAGCTCCATGCTATAGACCAAAATACCGTCTTGACG 747
|||||
Db 225 ThrArgArgValLeuAsnIleGlyProArgValProIleGlyProAsnProValIleIle 244
|||||
Qy 748 GGTCAAGACCCCCAACCCAA----- 768
|||||
Db 245 AspGlnLeuProProSerArgProValGlnIleMetLeuProArgProGlnProPro 264
|||||
Qy 769 GGACGAGCACCCTCTTAACATAACTTCTGGATCAGACCCCACTGAGCTTAACAGCAGC 828
|||||
Db 265 ProProGlyAlaIleSerIleValProGluThrAlaProProSerGlnGlnProGlyThr 284
|||||
Qy 829 ACTAAATGGGGGCAAACTTTTATGCTCATCCAGGGAGCTTTTCAAGCTCTTAACCTCC 888
|||||
Db 285 -----GlyAspArgLeuAsnLeuValAspGlyAlaTyrGlnAlaLeuAsnLeu 301
|||||
Qy 889 ACGACTCAGAGGCTACCTCTCTTGTGGTATGCTTAGCTTCGGGCCACCTTACTAT 948
|||||
Db 302 ThrSerProAspLysThrGlnGluCysTrpLeuCysLeuValAlaGluProProTyrTyr 321
|||||
Qy 949 GAAGGAATGGCTAGAGAGGGAATTCATCTGACAAAGAACAATAGACCAATGACACA 1008
|||||
Db 322 GluGlyValAlaValLeuGlyThrTyrSerAsnHisThrSerAlaProThrAsnCysSer 341
|||||
Qy 1009 TGGGGATCCCAAAATAAGCTTACCTTACTGAGGTTTCTGGAAAGGACCTTCGATAGGA 1068
|||||
Db 342 ValAlaSerGlnHisLysLeuThrLeuSerGluValThrGlyGlnGlyLeuCysIleGly 361
|||||
Qy 1069 AAGTTTCCCCATCCCAACAACACTTTGTAAACCACTGAAGCCTTTAATCAAACTCT 1128
|||||
Db 362 ThrValProLysThrHisGlnAlaLeuCysAsnThrThrLeuLysThrAsnLys----- 379
|||||
Qy 1129 GAGACTCAATATCTGGTACCTGGTTATGACAGGTGGTGGATGTAATACTGGATTAAACC 1188
|||||
Db 380 GlySerTyrTyrLeuValAlaProAlaGlyThrThrTrpAlaCysAsnThrGlyLeuThr 399
|||||
Qy 1189 CTTGTGTTTCCACCTGGTTTAAACCAAACTAAAGATTTTTCATTATGGTCCAAAT 1248
|||||
Db 400 ProCysLeuSerAlaThrValLeuAsnArgThrAspTyrCysValLeuValGluLeu 419
|||||
Qy 1249 GTTCCCGAGTGATTACTATCCCGAAAGCAATCCTTGATGAATATGACTACAGAAAT 1308
|||||
Db 420 TrpProArgValThrTyrHisProSerTyrValTyrSerGlnPheGlu---LysSer 438
|||||
```



```
Db 180 .....SerCysGluGlyLysCysAsnPro-----LeuValLeu 190
Qy 574 AGTTTCACTGAGAAAGAAAACAAGAAATATCTTAAATGCGTAAATGGTATGCTCTGG 633
Db 191 GlnPheThrGlnLysGlyArgGlnAlaSer-----TrpAspGlyProLysMetTrp 207
Qy 634 GGAATGGTATATATGAGGCTCGGGTAAACAACACGAGCTCCATTTCAACTATTTCGCTC 693
Db 208 GlyLeuArgLeuTyr--ArgThrGlyTyrAspProIleAlaLeuPheThrValSerArg 226
Qy 694 AAATAAACACGAGCTGGAGCTCCATGGCTATAGAGCAAAATAGGCTCTGACGGGTCAA 753
Db 227 GlnValSerThrIleThrProProGlnAlaMetGlyProAsnLeuValLeuProAspGln 246
Qy 754 AGACCCCAACCAAGGACCAAGCACCATCTCTAACATA-----ACT 795
Db 247 LysProProSerArgGlnSerGlnThrGlySerLysValAlaThrGlnArgProGlnThr 266
Qy 796 TCTGGATCAGACCCCACTGAGCTCTAACAGACGACTAAATATGGGGCAAAA-----846
Db 267 AsnGluSerAlaPro--ArgSerValAlaProThrThrMetGlyProLysArgIleGly 285
Qy 847 -----CTTTTAGCTCATCCAGGAGCTTTTCAAGCTCTTAAGCTCTTAAGCTCAAGCT 894
Db 286 ThrGlyAspArgLeuIleAsnLeuValGlnGlyThrTyrLeuAlaLeuAsnAlaThrAsp 305
Qy 895 CCAGAGGCTACCTCTCTTGTGCTATGCTTACGCTTCGGGCCACCTTACTATGAAGGA 954
Db 306 ProAsnLysThrLysAspCysThrPheCysLeuValSerArgProProTyrTyrGluGly 325
Qy 955 ATGGCTAGAGAGGGAATTCATGTGACAAAAGAACATAGAGACCAATGCACATGGGGA 1014
Db 326 IleAlaIleLeuGlyThrTyrSerAsnGlnThrAsnProProProSerCysLeuSerThr 345
Qy 1015 TCCCAAAATAGCTTACCTTACGTAGGTTCGGAAGGCAAGCCACCTGCATAGGAAGGTT 1074
Db 346 ProGlnHisLysLeuThrIleSerGluValSerGlyGlnGlyMetCysIleGlyThrVal 365
Qy 1075 CCCCACATCCCAACCAACCTTTGTAAACCAACACATCAAGCTTAAATCAACCTCGAGAGT 1134
Db 366 ProLysThrHisGlnAlaLeuCysAsnLysThrGlnGlnHis-----ThrGlyAla 383
Qy 1135 CAATATCTGGTACCTGGTTATGACAGGTGGTGGCATGTAACTAGGATTAAACCCCTTGT 1194
Db 384 HisTyrLeuAlaAlaProAsnGlyThrTyrTrpAlaCysAsnThrGlyLeuThrProCys 403
Qy 1195 GTTTCACCTTGGTTTTTAACCAAACTAAAGATTTTGCATTATGGTCCAAATGTTCCC 1254
Db 404 IleSerMetAlaValLeuAsnTrpThrSerAspPheCysValLeuIleGluLeuTrpPro 423
Qy 1255 CGAGTGTTACTATCCCGAAAAGCAATCTCTGATGAATATGACTACAGAAATCAT---1311
Db 424 ArgValThrTyrHisGln-----ProGluTyrValTyrThrHisPheAla 438
Qy 1312 -----CGACAAAAGAGAACCCATATCTCTGACACTTGCTGTGATGCTCGGA---1359
Db 439 LysAlaValArgPheArgArgGluProIleSerLeuThrValAlaLeuMetLeuGlyGly 458
Qy 1360 -----CTTGGAGTGGCAGAGGTGTAGGAACAGGAACAGCTGCCCTGGTCACGGA 1410
Db 459 LeuThrValGlyGlyIleAlaAlaGlyValGlyThrGlyThrLysAlaLeu-----475
Qy 1411 CCACAGCAGCTAGAAACAGGA--CTTAGTAACCTACATCGAATTTGTAACAGAGATCTC 1467
Db 476 -----LeuGluThrAlaGlnPheArgGlnLeuGlnMetAlaMetHisThrAspIle 492
Qy 1468 CAAGCCCTAGAAAATCTGTCAAGTAACCTGGAGGAATCCCTAACTCTCTTATCTGAAGTA 1527
Db 493 GlnAlaLeuGluSerIleSerAlaLeuGluLysSerLeuThrSerLeuSerGluVal 512
Qy 1528 GTCTCAGAAATAGAGGGGTATGATTATTATTTCTTAAAGAGGAGGATTATGTGTA 1587
|||||
```

Search completed: February 14, 2006, 16:07:34
Job time : 109.883 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 12:52:14 ; Search time 79.6345 Seconds
(without alignments)
3391.447 Million cell updates/sec

Title: US-10-723-552-3_COPY_5620_7533
Perfect score: 3468
Sequence: 1 ATGCATCCACGTTAAACCG.....CTAGCAGGGAAGCTGGCCGC 1914

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/absa/ABSSMBE spool/US10723552/runat 14022006 125141 12876/app query.fasta_1
-DB=UniProt -QPM=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USR=US10723552 @CGN 1.1 808 @runat 14022006 125141 12876 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot.sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3420	98.6	638	2	062707_PIG	062707 sus scrofa
2	3407	98.2	638	2	Q90LX4_GAMR	Q90LX4 porcine end
3	3402	98.1	638	2	062705_PIG	062705 sus scrofa
4	3402	98.1	638	2	Q90LX3_GAMR	Q90LX3 porcine end
5	3398	98.0	638	2	Q90LX5_GAMR	Q90LX5 porcine end
6	3363	97.0	638	2	Q5J1S9_PIG	Q5J1S9 sus scrofa
7	3293.5	95.0	641	2	Q8Q6Y9_GAMR	Q8Q6Y9 porcine end
8	3153.5	90.9	653	2	Q5QG07_GAMR	Q5QG07 porcine end
9	3150	90.8	642	2	Q5J1S8_SUSBA	Q5J1S8 sus barbatus
10	3147.5	90.4	653	2	Q8Q6Z0_GAMR	Q8Q6Z0 porcine end
11	3135.5	90.8	653	2	Q8Q6Z1_GAMR	Q8Q6Z1 porcine end
12	3018.5	87.0	659	2	Q8Q6V8_GAMR	Q8Q6V8 porcine end
13	2996.5	86.4	659	2	Q8Q6V6_GAMR	Q8Q6V6 porcine end
14	2995.5	86.4	659	2	Q8Q6V7_GAMR	Q8Q6V7 porcine end
15	2948.5	85.0	659	2	Q6UQ77_GAMR	Q6UQ77 porcine end
16	2943.5	84.9	659	2	Q4VFZ1_GAMR	Q4VFZ1 porcine end

17	2940.5	84.8	659	2	Q9Q9X3_GAMR	Q9Q9X3 porcine end
18	2928.5	84.4	659	2	Q6UQ78_GAMR	Q6UQ78 porcine end
19	2916.5	84.1	659	2	Q5J1S7_CETA	Q5J1S7 sus celeben
20	2916.5	84.1	659	2	Q6UQ79_GAMR	Q6UQ79 porcine end
21	2866.5	82.7	660	2	Q41172_GAMR	Q41172 porcine end
22	2862.5	82.5	660	2	Q8J4V7_GAMR	Q8J4V7 porcine end
23	2862.5	82.5	660	2	Q8Q6U6_GAMR	Q8Q6U6 porcine end
24	2862.5	82.5	2378	2	Q9Q1X4_GAMR	Q9Q1X4 porcine end
25	2857.5	82.4	660	2	Q6W4U7_PIG	Q6W4U7 sus scrofa
26	2852	82.2	654	2	Q6WGH9_PIG	Q6WGH9 sus scrofa
27	2851	82.2	654	2	Q90RL8_GAMR	Q90RL8 porcine end
28	2845	82.0	654	2	Q6W4U6_PIG	Q6W4U6 sus scrofa
29	2844.5	82.0	660	2	Q6W4V3_PIG	Q6W4V3 sus scrofa
30	2835.5	81.8	645	2	Q6W4U5_PIG	Q6W4U5 sus scrofa
31	2804.5	80.9	661	2	Q9IU66_GAMR	Q9IU66 porcine end
32	2788	80.4	655	2	Q516D0_GAMR	Q516D0 porcine end
33	2787	80.4	654	2	Q8UM98_GAMR	Q8UM98 porcine end
34	2786.5	80.3	649	2	Q6W4V2_PIG	Q6W4V2 sus scrofa
35	2769.5	79.9	660	2	Q8JEM7_GAMR	Q8JEM7 porcine end
36	2765	79.7	654	2	Q6W4V1_PIG	Q6W4V1 sus scrofa
37	2760.5	79.6	661	2	Q9IU77_GAMR	Q9IU77 porcine end
38	2685.5	77.4	612	2	Q6W4U8_PIG	Q6W4U8 sus scrofa
39	2683	77.4	633	2	Q6W4U9_PIG	Q6W4U9 sus scrofa
40	2642.5	76.2	613	2	Q9GLD5_PIG	Q9GLD5 porcine end
41	2628	75.8	589	2	Q90LX2_GAMR	Q90LX2 porcine end
42	2541.5	73.3	594	2	Q8WN83_PIG	Q8WN83 sus scrofa
43	2541.5	73.3	594	2	Q8WN84_PIG	Q8WN84 sus scrofa
44	2531.5	73.0	594	2	Q8WN77_PIG	Q8WN77 sus scrofa
45	2523.5	72.8	594	2	Q8WN86_PIG	Q8WN86 sus scrofa

ALIGNMENTS

RESULT 1
062707_PIG PRELIMINARY; PRT; 638 AA.
ID 062707_PIG
AC 062707;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Env protein [pig].
OS Sus scrofa [pig].
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=miniatore swine;
RX MEDLINE=98216827; PubMed=9557749;
RA Akiyoshi D.E., Denaro M., Zhu H., Greenstein J.L., Banerjee P., Fishman J.A.;
RT "Identification of a full-length cDNA for an endogenous retrovirus of miniatore swine.";
RL J. Virol. 72:4503-4507(1998).
DR EMBL; AF038600; AAC16765.1; -; mRNA.
DR HSP; P03385; IMOF.
DR SMR; 062707; 488-540.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env_polyprotein.
DR InterPro; IPR008981; FMuVrecept-bind.
DR Pfam; PF00429; TLV coat; 1.
SQ SEQUENCE 638 AA; 70997 MW; C8EC46D0D23BE4C8 CRC64;

Alignment Scores:
Pred. No.: 2,79e-268 Length: 638
Score: 3420.00 Matches: 637
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 98.6% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x 062707_PIG (1-638)

```
QY 1 ATGCATCCACGTTAAACGGGGCCACCTCCCGATTCCGGGTGGAAACCCGAAAGACTG 60
Db 1 MethHisProThrLeuAsnArgHisLeuProIleArgGlyGlyLeuProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCCTCCCATCGGTGGTTCCTTACTCTCTCAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGTTATGCGCATAGGACAGCCTGAACCTCCCATAAACCCCTATCTCTCACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAATGACTCCGGCAGACAGTATTAAATATCAACAACACTCAAGGGAGGCTCTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTATTATCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCTCCATCTCACGGATTATTATGTTGCCAGGACCAACCAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGGAATCCAGAGATTCTTTTGTGTAACAATGGAACGTGTAACTCTTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTGGAATGGCCAACTCTCAGCAGGATAGGTAAGTTTTCATTATGTCAC 480
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACCAAGCTCTGGCAATTTAATTAACCTGACCTGGATTAGAACTGGAGCCCCAAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCTTCAGACTAGATTACCTAAATAAATAGTTTTCATCTGAGAAAGGAAACAAGAA 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTTAAATGGTAAATGGTATGCTTGGGCAATGGTATATTATGGAGGCTCGGCT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyLysSerGly 220
QY 661 AAACAACCAAGGCTCCATCTCACTATTCGCTCAAAATAAACCACTGGAGCCTCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGGACCAATAACGCTGTGACGGGTCAAGACCCCAACCCCAAGGACCAAGACCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTCTGGATCAGACCCCTAGTCTTAACAGCACACTAAATAATGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerSerThrThrLysMetGly 280
QY 841 GCAAACTTTTTCAGCTCATCAGGGAGCTTTTCAAGCTCTTAACTCCACGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACTCTTCTTGTGTGGTATGCTTAGCTTCGGGCCCACTTACTTACTAAGAAATGGCT 960
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMetAla 320
QY 961 AGAAGAGGAAATTCATGTCGAAAGAACATAGACCAATGACATGCGGATGGGATCCCAA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpLysSerGln 340
QY 1021 AATAAGCTTACCTTACTGAGTTTCTGGAAAGGACCTGTCATAGGAAAGGTTCCCA 1080
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
```

```
QY 1081 TCCCAACCAACACCTTTGTAAACCACTGACGCTTTAATCAAACTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGTTACCTGCTTATGACAGGTGGTCATGTAATACTACTGATTAAACCCCTTGTGTTTCC 1200
Db 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTTTAACCAAACTAAAGATTTTTCGATTTATGGTCCAAATTTGTTCCCGAGTG 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAGCAATCCTTCATGAATATGACTACAGAAATCATCGCAAAAG 1320
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAACAACCATATCTCTGACACTTGTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTA 1380
Db 441 ArgGluProLysSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGAAACAGCTGCCCTGGTCACGGGACCAACAGCAGCTAGAAACAGGACTTAGTAAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTGTAAACAGAGATCTCCAAGCCCTAGAAAATCTGTCAGTAAACCTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACTCTTATCTGAAAGTAGTCTCAGCAATAGAAAGGAGGTTTAGATTATTA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGAGAGATTATGTAGCTTGAAGGAGGAATGCTGTTTTTATGTGGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
QY 1621 CATTCAGGGCCATCAGAGACTCCATGAACAGCTTAGAAGAAAGTTTCGAGAAGCGTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgGluLysArgArg 560
QY 1681 AGGAAAAGGAAACTACTCAAGGGTGGTTTCAGGGATGGTTCAACAGGCTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
QY 1741 GCTACCTTACTTTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTCTTACTACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATGTATTATTAAACAGTTAATTCCTTCATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgGileSerAlaValGln 620
QY 1861 ATCATGTTACTTAGACAACAGTACCAAGCCCGCTCTAGCAGGAAAGCTGGCGCG 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638
RESULT 2
Q90LX4_9GAMR PRELIMINARY; PRT; 638 AA.
ID Q90LX4_9GAMR PRELIMINARY; PRT; 638 AA.
AC Q90LX4_9GAMR PRELIMINARY; PRT; 638 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope protein.
OS Porcine endogenous type C retrovirus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=70540;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Blusch J.H., Seelmeier S., von der Helm K.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF402661; AAK94076.1; -; Genomic_DNA.
```


Qy	781	TCCTCTAACATAACTTCTCGATCAGACCCCACTCAGTCTAACACGACGACTAAATGGG	840
Db	261	SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly	280
Qy	841	GC AAAACTTTTACCTCATCCAGGGAGCTTTCAAGCTCTTAACTCCACGACTCCAGAG	900
Db	281	AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu	300
Qy	901	GCTACCTCTTCTTGTGTGGCTATGCTTACGCTTCGGGCCACCTTACTATGAAGGAATGGCT	960
Db	301	AlaThrSerSerCysThrLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAla	320
Qy	961	AGAAGAGGGAATTCATATGTGACAAAAGAACATAGAGACCAATGCACATGGGATCCCAA	1020
Db	321	ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTpeGlySerGln	340
Qy	1021	AATAAGCTTACCTTACTGAGGTTTCTGGAAAGGCACCTGCATAGGAAGGTTCCCCCA	1080
Db	341	AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro	360
Qy	1081	TCCCAACCAACACCTTGTAAACACACTGAAAGCCCTTTAATCAAACCTCTGAGAGCTCAATAT	1140
Db	361	SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr	380
Qy	1141	CTGGTACCTGGTTATGACAGGTGGTGGCATGTAATATCGATTAACCCCTTGTGTTC	1200
Db	381	LeuValProGlyTyrAspArgTptTpeAlaCysAsnThrGlyLeuThrProCysValSer	400
Qy	1201	ACCTTGGTTTTTAAACCAACTAAAGATTTTGTGATTTATGGTCCAAATGTTTCCCGAGTG	1260
Db	401	ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal	420
Qy	1261	TATTACTATCCCAAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCCACAAAAG	1320
Db	421	TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys	440
Qy	1321	AGAAACCCCATPCTCTGACACTGCTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGA	1380
Db	441	ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal	460
Qy	1381	GGACAGGACAGCTGCCCTGGTCACGGNACCACACAGCTAGAAAACAGCACTTAGTAAC	1440
Db	461	GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn	480
Qy	1441	CTACATCGAATTGTAACAGAGATCTCCAAGCCCTAGAAAAATCTGTCAGTACCTGGAG	1500
Db	481	LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu	500
Qy	1501	GAATCCCTAACCTCTTATCTGAAGTAGTCTTACAGAAATAGAGAGGGTTAGATTTATTA	1560
Db	501	GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu	520
Qy	1561	TTTCTAAAAGAGGAGGATATGTGTAGCTTGAAGGAGAAATGCTGTTTTTATGTGGAT	1620
Db	521	PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp	540
Qy	1621	CATTACAGGSGCCATCAGACACTCCATGAACAAGCTTAGAGAAGGTTGAGAGAGCGTCGA	1680
Db	541	HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg	560
Qy	1681	AGGGAAGGAAACTACTCAAGGTTGTTGAGGATGTTTCAACAGGCTCTCTTTGGTTG	1740
Db	561	ArgGluLysGluThrThrGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTpeLeu	580
Qy	1741	GCTACCTTACTTTCGCTTTAACAGACCCCTTAATAGTCCTCTCCTCTGTGTACTCACAGTT	1800
Db	581	AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal	600
Qy	1801	GGGCCCATGTATTATTAAAGTTAATGCTTCATTAGAGACGGAATAGTCAGTCCAG	1860
Db	601	GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln	620
Qy	1861	ATCATGTGACTTAGACAACAGTACCAAGCCCGTCTTAGCAGGGAAGCTGGCCGC	1914

301	Ala	Thr	Ser	Cys	I	Pro	Leu	Ala	Leu	Gly	Pro	Pro	Tyr	Tyr	Gly	Gly	Met	Ala	320
961	AGA	AGG	GGG	AAA	T	T	CA	T	G	T	G	C	A	A	A	A	A	A	1020
321	Arg	Arg	Gly	Lys	Phe	Asn	Val	Thr	Lys	Glu	His	Arg	Asp	Gln	Cys	Thr	Trp	340	
1021	A	A	T	A	A	G	C	T	T	A	C	T	T	A	C	T	T	1080	
341	Asn	Lys	Leu	Thr	Leu	Thr	Glu	Val	Ser	Gly	Lys	Gly	Thr	Cys	Ile	Gly	Lys	360	
1081	T	C	C	C	A	C	A	C	A	C	T	T	T	G	A	A	C	1140	
361	Ser	His	Gln	His	Leu	Cys	Asn	His	Thr	Glu	Ala	Phe	Asn	Gln	Thr	Ser	Glu	380	
1141	C	T	G	T	A	C	T	T	A	C	A	G	G	T	G	C	G	1200	
381	Leu	Val	Pro	Gly	Tyr	Asp	Arg	Tyr	Trp	Trp	Ala	Cys	Asn	Thr	Gly	Leu	Thr	400	
1201	A	C	T	T	G	T	T	T	T	A	C	A	A	A	T	T	T	1260	
401	Thr	Leu	Val	Phe	Asn	Gln	Thr	Lys	Asp	Phe	Cys	Ile	Met	Val	Gln	Ile	Val	420	
1261	T	A	T	A	C	T	A	T	C	C	C	A	A	A	A	A	A	1320	
421	Tyr	Tyr	Tyr	Pro	Glu	Lys	Ala	Ile	Leu	Asp	Glu	Tyr	Asp	Tyr	Arg	Asn	His	440	
1321	A	G	A	A	C	C	A	T	A	C	T	C	T	G	T	G	T	1380	
441	Arg	Glu	Pro	Ile	Ser	Leu	Thr	Leu	Ala	Val	Met	Leu	Gly	Leu	Gly	Val	Ala	460	
1381	G	A	A	C	A	G	G	A	C	A	C	T	G	C	T	T	G	1440	
461	Gly	Thr	Gly	Thr	Ala	Ala	Leu	Val	Thr	Gly	Pro	Gln	Gln	Leu	Glu	Thr	Gly	480	
1441	C	T	A	C	A	T	G	A	A	G	A	T	C	T	C	A	A	1500	
481	Leu	His	Arg	Ile	Val	Thr	Glu	Asp	Leu	Gln	Ala	Leu	Glu	Lys	Ser	Val	Ser	500	
1501	G	A	A	T	C	C	T	A	C	T	T	A	C	T	G	A	G	1560	
501	Glu	Ser	Leu	Thr	Ser	Leu	Ser	Glu	Val	Val	Leu	Gln	Asn	Arg	Arg	Gly	Leu	520	
1561	T	T	T	C	A	A	A	G	A	G	A	T	T	A	C	T	G	1620	
521	Phe	Leu	Lys	Glu	Gly	Lys	Cys	Val	Ala	Leu	Lys	Glu	Glu	Cys	Phe	Tyr	Val	540	
1621	C	A	T	T	C	A	G	G	G	C	C	A	T	C	A	A	C	1680	
541	His	Ser	Gly	Ala	Ile	Arg	Asp	Ser	Met	Asn	Lys	Leu	Arg	Glu	Arg	Lys	Arg	560	
1681	A	G	G	A	A	A	G	A	A	C	T	A	C	T	T	G	A	1740	
561	Arg	Glu	Lys	Glu	Thr	Thr	Gln	Gly	Tyr	Phe	Glu	Gly	Tyr	Phe	Asn	Arg	Ser	580	
1741	G	C	T	A	C	C	T	A	C	T	T	A	C	A	G	A	C	1800	
581	Ala	Thr	Leu	Leu	Ser	Ala	Leu	Thr	Gly	Pro	Leu	Ile	Val	Leu	Leu	Leu	Thr	600	
1801	G	G	S	C	A	T	G	T	A	A	C	A	A	G	T	T	A	1860	
601	Gly	Pro	Cys	Ile	Ile	Asn	Lys	Leu	Ile	Ala	Phe	Ile	Arg	Glu	Arg	Ile	Ser	620	
1861	A	T	C	A	T	G	T	A	C	T	A	G	A	A	C	C	G	1914	

341	AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyValPro	360
1081	TCCACCACACACCTTGTAAACACACTGAAAGCCCTTAAATCAAACCTCTGAGAGCTCAATAT	1140
361	SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr	380
1141	CTGGTACCTGGTTATGACACAGGTGGTGGCATGTAATACCTGGATTAACCCCTTCTGTTTCC	1200
381	LeuValProGlyTyrAspArgTTPtAlaCysAsnThrGlyLeuThrProCysValSer	400
1201	ACCTTGGTTTTTAAACCAACTAAAGATTTTTCGATTATGGTCCAAATGTTTCCCGAGTGG	1260
401	ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal	420
1261	TATTACTATCCCGAAAAAGCAATCCTTGTGATGAATATGACTACAGAAATCATCCACAAAAG	1320
421	TyrtYrTyrProGlnLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys	440
1321	AGAAACCCCATPCTCTGCACACTGCTGTGATGCTCGGACTTGGAGTGGCGAGGTGTA	1380
441	ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal	460
1381	GGACACAGGACAGCTGCCCTGGTCACGGGACACACAGCAGCTAGAAAACAGGACTTAGTAAC	1440
461	GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn	480
1441	CTACATCGAATTGTAACAGAGATCTCCAAGCCCTAGAAAAATCTGTCAGTACACCTGGAG	1500
481	LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu	500
1501	GAATCCCTAACCTCTTATCTGAAGTAGTCTTACAGATAGAGAGGGTTAGATTTATTATA	1560
501	GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu	520
1561	TTTCTAAAAGAGGAGGATTATGTGTAGCTTGAAGGAGGAATGCTGTTTTTATGTGGAT	1620
521	PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp	540
1621	CATTCAAGGCGCCATCAGACACTCCATGAACAAGCTTAGAGAAGGTTTGAGAGAGCTCGA	1680
541	HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgIleuArgLeuGluLysArgArg	560
1681	AGGGAAAAGGAACTACTCAAGGGTGGTTTGAGGATGTTTCAACAGGCTCTCTTTGGTTG	1740
561	ArgGluLysGluThrThrGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrLeu	580
1741	GCTACCCCTACTTCTGCTTTACAGGACCCCTTAATAGTCCTCTCTCTGTTTACTCACAGTT	1800
581	AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal	600
1801	GGGCCCATGTATTATTAAACAGTTAATGCTCTTCATTAGAGACGGAATAGTCAGTCCAG	1860
601	GlyProCysIleIleLeasnlLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln	620
1861	ATCATGTGCTACTTAGACAACAGTACCAAGCCCGTCTTAGCAGGGAAGCTGGCCGC	1914

1081	QY	TCCACACACACCTTTGTAAACCACTCGAAGCCCTTAAATCAAAACCTCTGAGACTCAATAT	
361	Db	SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr	380
1141	QY	CTGGTACTCGTTATGACAGCTGGTGGGCATGTAATACTGGATTAAACCCCTGTGTGTTCC	1200
381	Db	LeuValProGlyTyrAspArgTrpTyrAlaCysAsnThrGlyLeuThrProCysValSer	400
1201	QY	ACCTGTGGTTTTAAACCAACTAAAGATTTTTTGCATTATGCTCCAAATGTTTCCCCGAGTG	1260
401	Db	ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal	420
1261	QY	TATTACTATCCCGAAAAAGCAATCTCTTGATGAATATGACTACAGAAATCATCGACAAAAG	1320
421	Db	TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys	440
1321	QY	AGAAACCCCATCTCTGACACTGCTGCTGATGCTCGACTCGGACTCGGAGTGGCAGCGTGTA	1380
441	Db	ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal	460
1381	QY	GGAAACAGGAACAGCTCCCTCGTGCAGGACACACACAGCTAGAAAACAGACATTAGTAAC	1440
461	Db	GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn	480
1441	QY	CTACATCGAATTGTAAACAGAAGATCTCCAAGCCCTAGAAAAATCTCTCAGTAACTCGGAG	1500
481	Db	LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu	500
1501	QY	GAATCCCTAACCTCTTATCTGAAGTAGTCTTACAGAATAGAAAGGGTTAGATTATTATTA	1560
501	Db	GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu	520
1561	QY	TTTCTAAAGAGGAGGATTATGTGTAGCTTGAAGGAGGAATGCTGTTTTTATGTCGAT	1620
521	Db	PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp	540
1621	QY	CATTCAAGGCGCAATCAGACACTCCATGAACAAGCTTAGAGAAAGGTTTGAGAAAGCCTCGA	1680
541	Db	HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuLysArgArg	560
1681	QY	AGGGAAAGGAACTACTCAAGGGTGGTTTGAGGGATGCTTCAACAGGCTCTCTTTGGTTG	1740
561	Db	ArgGluLysGluThrThrGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrLeu	580
1741	QY	GCTACCTTACTTTCTGCTTTAACAGACCCCTTAATAGTCCTCTCTCTGTTACTCAGATT	1800
581	Db	AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrVal	600
1801	QY	GGGCCCATGTATTATAACAAGTTAATTCGCTTCATTAGAGACGAATAAGTCAGTCCAG	1860
601	Db	GlyProCysIleIleLeasnlLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln	620
1861	QY	ATCATGTGTTACTTAGACAACAGTACCACCAAGCCGCTTAGCAGGGAAGCTGGCCGC	1914

```
Db 621 ileMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638
|||||
RESULT 3
O62705_PIG
ID O62705_PIG PRELIMINARY; PRT; 638 AA.
AC O62705
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Env protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sub.
NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98216827; PubMed=9557749;
RA Akiyoshi D.E.; Denaro M., Zhu H., Greenstein J.L., Banerjee P.,
RA Fishman J.A.;
RT "Identification of a full-length cDNA for an endogenous retrovirus of
RT miniature swine.";
RL J. Virol. 72:4503-4507(1998).
DR EMBL; AF038599; AAC16762.1; -; mRNA.
DR HSP; P03385; IMOF.
DR SMR; O62705; 488-540.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env polyprotein.
DR InterPro; IPR008981; FmuIVrecept-bind.
DR Pfam; PF00429; TLV_coat; 1.
SQ SEQUENCE 638 AA; 70980 MW; CFC394814FD23FEA CRC64;

Alignment Scores:
Pred. No.: 8,04e-267 Length: 638
Score: 3402.00 Matches: 634
Percent Similarity: 99.7% Conservative: 2
Best Local Similarity: 99.4% Mismatches: 2
Query Match: 98.1% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x O62705_PIG (1-638)
QY 1 ATGCATCCACGCTTAAACGGGGCCACCTCCGATTCGGGTGGGAAACCGAAAGACTG 60
Db |||||::|||
1 MethHisProThrLeuSerArgHisLeuProIleArgGlyGlyLeuProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCTTACTCTCTCAATAACCTCTCAG 120
Db |||||
21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGGTATGCGCATAGGACAGCCTGAACCTCCCATAAACCTTATCTCTCACCTGG 180
Db |||||
41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAAGTACCGGCACAGGATTAATATCAACACACTCAAGGGGAGGCTCTTTA 240
Db |||||
61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGCGCTGATCTATACCTTTGCTTCAGATCAGTATTTCCTAGTCTGACCTCA 300
Db |||||
81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCCGATATCTCCATGCTCAGCGATTTTATGTTTGGCCAGGACCAACCAATAATGGA 360
Db |||||
101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGGAAATCCAGAGATTTCTTTGTAAACCAATGAACTGTGTAACCTCTAAT 420
Db |||||
121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTGGAAATGGCCAACTCTCAGCAGGATAGGTAAGTATTTCTTATGTCAC 480
```

```
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyr-ValAsn 160
|||||
481 ACCTATACCGAGCTCTGGACAAATTAATTAACCTGAGCTGGATTAGAACTGGAGCCCAAG 540
Db |||||
161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCCTCTCTTCAGACCTAGATTACTAAATAATAGTTTCACTGAGAAAGGAAACAAGAA 600
Db |||||
181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyGlnGlu 200
QY 601 AATATCTCTAAATGGTAAATGGTATGTTCTTGGGGAATGGTATATTATGGAGGCTCGGGT 660
Db |||||
201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGly 220
QY 661 AAACAACAGGCTCCATCTTAATCTATTCGCTCAAAATAAACACAGCTGGAGCCTCCCAATG 720
Db |||||
221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGGACCAATACGGTCTTCACGGGTCAAGACCCCAACCCCAAGGACCAAGACCA 780
Db |||||
241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTTCTGGATCAGACCCCACTGAGTCTAAACAGCACACTAAATGGGG 840
Db |||||
261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerSerThrThrLysMetGly 280
QY 841 GCAAAACCTTTTACCTCATCATCAGGAGCTTTTCAAGCTCTTAATCTCACAGCTCCAGAG 900
Db |||||
281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTTCTTGGTATGCTAGCTTCGGGGCCACCTTACTATGAAGGATGGCT 960
Db |||||
301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGlnGlyMetAla 320
QY 961 AGAAGAGGAAATTCATATGTGACAAAGAAATAGAGACCAATGACATGGGGATCCCAA 1020
Db |||||
321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTAGGTTTCTGGAAAGGACCTGCATAGGAAAGGTTCCTCCCA 1080
Db |||||
341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACACACCTTTGTAAACCACTGAACCTTTAATCAAACTCTGAGAGTCAATAT 1140
Db |||||
361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGGTACCTGTTATGACAGGTGGTGGGCATGTAAATACTGGATTAAACCCCTTGTGTTCC 1200
Db |||||
381 LeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTAAACCAAACTAAAGATTTTTCATATGCTTCGATTCGATTCCTCCAGAGT 1260
Db |||||
401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAGCAATCTTGATCAATATGACTACAGAAATCATCGACAAAG 1320
Db |||||
421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAGAACCATATCTCTGACACTTGTGTGTGTGCTCGGACTTGGAGTGGCAGCAGGTGTA 1380
Db |||||
441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAAACAGAAACAGCTGCCCTGGTCCAGGGACACAGCAGCTAGAAAAACAGGACTTAGTAAC 1440
Db |||||
461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTGTAACAGAGAGATCTCAAGCCCTAGAAAAATCTGTCAAGTAACCTGAG 1500
Db |||||
481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACCTCTTATCTCTGAAGTAGTCTCTACAGATAGAGAGGCTTAGATTATTA 1560
Db |||||
```

Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
 QY 1561 TTTCTAAAGAGGAGGATTATGTGTACCTTGAAGGAGGAATGCTGTTTTATGTGGAT 1620
 Db 521 PheLeuLysGluGlyLeuCysValAlaLeuLysGluGlyCysPheTyrValAsp 540
 QY 1621 CATTCAGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAAAGTTGGAGAAGCGTCGA 1680
 Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
 QY 1681 AGGAAAAAGAAACTACTCAAGGGTGGTTTCAGGGATGTTCAACAGGTCTCTTTGGTTG 1740
 Db 561 ArgGluLysGluThrThrGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrLeu 580
 QY 1741 GCTACCTTACTTCTGCTTTTAAAGACCTTAATAGTCTCTCTCTCTCTCTCTCTCTCT 1800
 Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
 QY 1801 GGGCCATGTATTATTAAACAGTAAATGCTTCAATAGAGAACGAATAAGTGCAGTCCAG 1860
 Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
 QY 1861 ATCATGTACTTAGACACAGTACCMAAGCCCGTCTAGCAGGGAAGCTGGCCGC 1914
 Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 4

Q90LX3_9GAMR
 ID Q90LX3_9GAMR PRELIMINARY; PRT; 638 AA.
 AC Q90LX3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Envelope protein.
 OS Porcine endogenous type C retrovirus.
 OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
 OX NCBI_TaxID=70540;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Blusch J.H., Seelmeier S., von der Helm K.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF402662; AAK94077.1; -, Genomic_DNA.
 DR HSP; P03385; IMOF.
 DR SMR; Q90LX3; 488-540.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR002050; Env_polyprotein.
 DR InterPro; IPR008981; FmLVrecept-bind.
 DR Pfam; PF00429; TLV_coat; 1.
 KW Envelope protein.
 SQ SEQUENCE 638 AA; 70980 MW; CFC394814FD23FEA CRC64;

Alignment Scores:

Pred. No.: 8,04e-267 Length: 638
 Score: 3402.00 Matches: 634
 Percent Similarity: 99.7% Conservative: 2
 Best Local Similarity: 99.4% Mismatches: 2
 Query Match: 98.1% Indels: 0
 DB: 2 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x Q90LX3_9GAMR (1-638)

QY 1 ATGCATCCACAGTTTAAACCGCGCCACCTCCCGATTCGGGGTGGAAAGCCGAAAGACTG 60
 Db 1 MechisProThrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
 QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG 120
 Db 21 LysIleProLeuSerPheAlaSerIleAlaTyrPheLeuThrLeuSerIleThrSerGln 40
 QY 121 ACTAATGGTATGGCATAGAGACAGCTCGAATCTCCATAAACCTTATCTCACTCG 180

Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
 QY 181 TTAATTACTGACTCCGGCACAGGTATTAAATCAACACACTCAAGGGGAGGCTCCTTTA 240
 Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
 QY 241 GGAACTGTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTATTCCTAGTCTGACCTCA 300
 Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
 QY 301 CCCCAGATATCTCCATGCTCAGGATTTTATCTTTCGCCAGGACCCACCAATAATCGA 360
 Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
 QY 361 AAAATTTCCGGAATCCAGAGATTTCTTTTGTAAACAATGAACTGTGTAACTCTAAT 420
 Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
 QY 421 GATGGATATTGGAAATGGCCAACTCTCAGCAGATAGGGTAAGTTTCTTATGTCAAC 480
 Db 141 AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn 160
 QY 481 ACCTATACAGCTCTCGACAAATTTAATCTGACCTGGATTAGAACTGGAGCCCCAAG 540
 Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
 QY 541 TGCTCTCTTTCAGACCTAGATTACTAAAATAAGTTTCACTGAGAAGGAAAAACAAGAA 600
 Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
 QY 601 AATATCTTAAATGGGTAAATGGTATCTCTTGGGAAATGGTATATTATGGAGGCTCGGGT 660
 Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlySerGly 220
 QY 661 AAACAACAGCTCTCCATCTTAACCTATTCGGCTCAAAATAAACAGCTGGAGCTCCAAATG 720
 Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
 QY 721 GCTATAGGCAAAATACGGTCTTGACGGGTCAAGAGCCCCCAACCCCAAGGACAGGACCA 780
 Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
 QY 781 TCCTCTAACATAACTCTCGGATCAGACCCCACTGAGTCTAACACGACGACTTAAATGGG 840
 Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerThrThrLysMetGly 280
 QY 841 GCAAAACCTTTTAGCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAG 900
 Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
 QY 901 GCTACCTCTTCTTGTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGGAATGGCT 960
 Db 301 AlaThrSerSerCysTrpLeuLysLeuAlaLeuGlyProProTyrTyrGlyLysMetAla 320
 QY 961 AGAAGAGGGAATTCATATGTGACAAAAGAACATAGAGACCAATGCATATGGGATCCCAA 1020
 Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
 QY 1021 AATAAGCTTACCTTACTGAGGTTCTGGAAAAGGACCTGCGATAGAAAAGTTTCCCCCA 1080
 Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
 QY 1081 TCCCACCAACACCTTTGTAAACACACTGAGCCCTTTAATCAAACTCTGAGAGTCAATAT 1140
 Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
 QY 1141 CTGGTACCTGGTTATCAGAGGTGGTGGCATGTAATCTGGAATAACCCCTTGTGTTTC 1200
 Db 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
 QY 1201 ACCTTGGTTTAAACCAAACTAAAGATTTTTCATTATGGTCCAAATGTTTCCCCGAGTG 1260
 Db 401 ThrLeuValPheAsnGlnThrLysArgPheCysIleMetValGlnIleValProArgVal 420

```
QY 1261 TATTACTATCCGAAAAAGCAATCCTTGATCAATATGACTACAGAAATCATCGACAAAAG 1320
Db 421 TyrtYrtyrProGluLysAlaIleuAspGluYrAspTyrArgAsnHlaRglnLys 440
QY 1321 AGAGAACCCATATCTCTGACACTTGCTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTA 1380
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGGAACAGCTGCCCTGTCACGGGACACACAGCAGCTAGAAACAGGACTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTGTAAACAGAGATCTCCAAGCCCTAGAAAAATCTGTCAGTAACCTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAAATCCCTAAACCTCTTATCTCAAGTAGTCTCTACAGATAGAGAGGGTTAGATTTATTA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnA-gA-gGlyLeuAspLeuLeu 520
QY 1561 TTCTTAAAGAGGAGGATTAATGTAGCTTGAAGGAGGAATGCTGTTTTTATTTGTGGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTy-rValAasp 540
QY 1621 CATTGAGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAAGGTTGGAGAGCGTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgarg 560
QY 1681 AGGGAAAAAGGAAACTACTCAAGGGTGGTTTGAGGGATGGTTCAACAGAGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnA-gSerProTrpLeu 580
QY 1741 GCTACCCCTACTTCTGCTTTAACAGGACCCCTTAATAGTCCTCTCTCTGTTACTCACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATCTATTATTAACAGTTAATTCCTTCAATTAGAGAACGTAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgGluLeuAlaValGln 620
QY 1861 ATCATGGTACTTAGAACACAGTACCAGGCCGCTAGCAGGGAAGCTGGCCGC 1914
Db 621 IleMetValLeuArgGlnGlnTy-rGlnSerProSerSerA-gGluAlaGlyArg 638

RESULT 5
Q90LX5_9GAMR
ID Q90LX5_9GAMR PRELIMINARY; PRT; 638 AA.
AC Q90LX5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope protein.
OS Porcine endogenous type C retrovirus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=70540;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Blusch J.H., Seelmeier S., von der Helm K.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF402660; AAK94075.1; -; Genomic_DNA.
DR HSP; P03385; 1MOP.
DR SMR; Q90LX5; 488-540.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env polypeptide.
DR InterPro; IPR008981; FMuLVrecept-bind.
DR Pfam; PF00429; TUV_coat; 1.
KW Envelope protein.
SQ SEQUENCE 638 AA; 71053 MW; 15B3A4BCDA40B3C7 CRC64;
Alignment Scores:
```

```
Pred. No.: 1.7e-266 Length: 638
Score: 3398.00 Matches: 633
Percent Similarity: 99.4% Conservative: 1
Best Local Similarity: 99.2% Mismatches: 4
Query Match: 98.0% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_copy_5620_7533 (1-1914) x Q90LX5_9GAMR (1-638)

QY 1 ATGCATCCCAAGTTAAACCGCGCCACCTCCGATTCGGGGTGAAGCCGAAAGACTG 60
Db 1 MethisProthrLeuSerA-gArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAAGCTTCGCCTCCATCGCGTGGTTCCTACTCTGTCAATAACCTCTCAG 120
Db 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGGTATGCGCATAGGAGACAGCTGAACCTCCATAAACCTTATCTCTCACCTGG 180
Db 41 ThrAsnGlyMetA-gIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGCGACAGGTATTAAATATCAACACTCAAGGGAGGCTCCTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGCTCAGATCAGTATTCTTAGTCTCACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTy-rValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCTCCATGCTCACGATTTTATGTTGCCAGGACCAACCAAAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTy-rValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGAATCCAGAGATTTCTTTTCTTAAACAATGGAACTGTGTAACCTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATCGATATTGGAATGGCAACCTCTCAGCAGATAGGGTAAGTTTTTCTTAGTCTCAAC 480
Db 141 AspGlyTy-rTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTy-rValAsn 160
QY 481 ACCTATACCACTCTGGACAAATTTAATTAACCTGATTAAGATAGAACCTGGAGCCCAAG 540
Db 161 ThrTy-rThrSerSerGlyGlnPheAsnTy-rLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCCTTCAGACCTAGATTACTTAAATAAGTTTCACTGAGAAAGGAAAACAGAA 600
Db 181 CysSerProSerAspLeuAspTy-rLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTTAAATGGTAAATGGTATGCTTGGGGAATGGTATATTATGGAGGCTCGGGT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTy-rGlyGlySerGly 220
QY 661 AAACAACCCAGCTCCATTTCAACTATTTCGCTCAAAATAAACACCACTGGAGCCCTCAAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuLeuProProMet 240
QY 721 GCTATAGACCAATAATCGTCTTGACGGTCAAGACCCCAACCCCAAGGACAGGACCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTTCTGGATCAGCCCCACTCAGTCTAACAGCACCACCTAAATAATGGG 840
Db 261 SerCysAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
QY 841 GCAAAACCTTTTATAGCCTCATCAGGAGGCTTTTCAAGCTCTTAACTCCACGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTTCTGTTGGCTATGCTTAGCTTCGGGGCCACCTTACTATGAGAGGATGGCT 960
Db 301 ThrThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTy-rTy-rGlyMetAla 320
```

```
QY 961 AGAAGAGGAAATTCATGTGACAAAGAACATAGAGACCATATGCATATGGGATCCAA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTCGAAAGGCACCTGCATAGGAAAGTCCCCCA 1080
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACCAACCTTTGTAACCACTGAGCCTTTAATCAAACTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTrp 380
QY 1141 CTGGTACCTGGTTATGACAGGTGGTGGCATGTAATCTGATTAACCCCTTGTGTTCC 1200
Db 381 LeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTGTTTTAAACCAACTAAAGATTTTTCATTATGTCCTCAAAATGTTCCCGAGT 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAG 1320
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAGAACCATATCTCTGACACTTCTGTGATGTCGAGCTTGAGTGGCAGCGGTGTA 1380
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGACAGGAACAGCTGCCCTGTGTCACGGGACACAGCAGCTAGAAACAGGACTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTTGTAACAGAAAGATCTCCAAGCCCTAGAAAAATCTGTCAAGTAACCTGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACCTCTTATCTGAAGTAGTCTCTACAGATAGAGAGGGTTAGATTATA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGATTATGTAGCTTGAAGGAGGAATCTGTGTTTTATGTGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
QY 1621 CATTCAGGGCCATCAGAGACTCCATGAACAAAGCTTAGAGAAAGGTTGGAGAACGTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGAAAGGAACTACTCAAGGTGGTTGAGGGATGGTTCAACAGGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
QY 1741 GCTACCTTACTTCTGCTTTAACAGGACCTTATAGTCTCTCTCTCTCTCTCTCTCTCT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATGTATTATTAAACAGTTAATTCCTTCAATTAGAACCAATAGTGCAAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGTACTTAGACACAGTACCAAGCCCGCTCTAGCAGGGAAGCTGGCCGC 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638
```

RESULT 6

```
Q5J1S9_PIG
ID Q5J1S9_PIG PRELIMINARY; PRT; 638 AA.
AC Q5J1S9;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;
```

```
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15596862; DOI=10.1128/JVI.79.1.649-654.2005;
RA Niebert M., Tonjes R.R.;
RT "Evolutionary spread and recombination of porcine endogenous
RT retroviruses in the suiformes.";
RL J. Virol. 79:649-654(2005).
DR EMBL; AY534304; AAT06256.1; -; mRNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env_polyprotein.
DR Pfam; PF00429; tlv_coat; 1.
KW Envelope protein.
SQ SEQUENCE 638 AA; 71021 MW; 9428D75C6F21B3B8 CRC64;
```

Alignment Scores:

Pred. No.:	1.17e-263	Length:	638
Score:	3363.00	Matches:	627
Percent Similarity:	98.7%	Conservative:	3
Best Local Similarity:	98.3%	Mismatches:	8
Query Match:	97.0%	Indels:	0
DB:	2	Gaps:	0

US-10-723-552-3_COPY_5620_7533 (1-1914) x Q5J1S9_PIG (1-638)

```
QY 1 ATGCATCCACGTTAAACCCGCGCCACCTCCCGATTCCGGGTGAAAGCCGAAAGACTG 60
Db 1 MethiSprothrLeuAsnArgArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATAACTCTCAG 120
Db 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATCGTATCGCATAGGACAGCCTGAACTCCCAATCAACCTTATCTCTCACTCG 180
Db 41 ThrAsnGlyMetArgIleGlyAsnSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGCGCACAGGTATTAAATATCAACAACTCAAGGGAGGCTCTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGCTCAGATCAGTTATCTTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuAlaThr 100
QY 301 CCCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCCACCAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGlu 120
QY 361 AAACATTGCGGAATCCAGAGATTTCTTTGTAAACAATGGAAGTGTAACTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGGATATTGGAATGGCCAACTCTCAGCAGGATAGGGTAAGTTTTTCTTATGTCAAC 480
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACCACTCTGGACAAATTTAATTAATCTGACCTGGATTAGAACTGGAGCCCCAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCTCTTTCAGACCTTAGATTACCTAAAAATAAGTTTCACTGAGAAGGAAACAAAG 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTAAAAATGGGTAATGGTATCTCTTGGGAATGGTATATTATGAGGCTCGCGT 660
```

```
Db      201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTy-TyrGlySerGly 220
QY      661 AAACACAGCGCTCCATTCTAACTATTTCGCCTCAAAATAAACACGCTGGAGCCTCCAATG 720
Db      221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY      721 GCTATAGACCAAAATACCGCTTGTAGCGGTCAAGACCCCAACCAAGGACCAAGACCA 780
Db      241 AlaThrGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY      781 TCCTCTAAATAAATCTTCGGATCAGACCCCACTAGTCTAAACAGCACCACTAAATAATGGG 840
Db      261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThryLysMetGly 280
QY      841 GCAAACTTTTATAGCTCATCCAGGAGCTTTTCAAGCTCTTAAGCTCTTAATCTCCAGACTCCAGAG 900
Db      281 AlalysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY      901 GCTACCTCTTCTGTGTGGCTATGCTTAGCTTCGGGCCCACTTACTACTATGAAGGAATGGCT 960
Db      301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyTrpGluGlyMetAla 320
QY      961 AGAAGAGGAAATTCATGTGCAAAAGAACAATAGAGACCAATGCACATGGGGATCCCAA 1020
Db      321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY      1021 AATAAGCTTACCTTACTAGGTTTCTCGAAGGACACCTGCATAGGAAGGTTCCTCCCA 1080
Db      341 AsnLysLeuThrLeuThrGluValPheGlyLysGlyThrCysIleGlyLysValProPro 360
QY      1081 TCCCAACCAACACTTTGTAAACACACTGAAGCTTTTAACTCAAACTCTGAGAGTCAATAT 1140
Db      361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTy 380
QY      1141 CTGTGTACCTGGTTATGACAGGTGGTGGCATGTAATACTGGAATTAACCCCTTGTGTTC 1200
Db      381 LeuValProGlyTyTrpAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY      1201 ACCTTGGTTTAAACAACTAAAGATTTTGCATTATGGTCCAAATTTGTTCCCGAGTG 1260
Db      401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY      1261 TATTACTATCCGAAAGAACCAATCTTGATGAATATGACTACAGAAATCATCGACAAAAG 1320
Db      421 TyTrpTyTrpProGluLysAlaIleLeuAspGluTyTrpAspTyTrpGlyAsnHisArgGlnLys 440
QY      1321 AGAGAACCCATATCTCTGACACTTCTGTGATGCTCGGACTTGGAGTGGCGACAGGTGTA 1380
Db      441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY      1381 GGAAACAGGAACAGCTGCCCTGTCTACGGGACACAGCAGCTAGAAACAGGACTTAGTAAC 1440
Db      461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY      1441 CTACATCGAATGTAAACAGACATCTCCAGGCCCTAGAAATCTGCTAGTAACCTCGAG 1500
Db      481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY      1501 GAATCCCTAAACCTCTTATCTCAAGTAGTCTTACAGAATAGAGAAGGGTTAGATTATTA 1560
Db      501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY      1561 TTTCTAAAGAGGAGGATATGTGTAGCTTGAAGAGGAATGCTGTTTTTATGTGGAT 1620
Db      521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyValAsp 540
QY      1621 CATTCAGGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAGGTGTGAGAAGCGTGA 1680
Db      541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgGluLysArgArg 560
QY      1681 AGGGAAAAGGAACTACTCAAGGGTGGTTTGAAGGATGTTCAACAGGTCTCTTTGGTTG 1740
Db      561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
```

```
QY      1741 GCTACCCCTACTTTCTGCTTTTAAACAGGACCCCTTAATAGTCTCTCTCTTACTCAGTT 1800
Db      581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY      1801 GGGCCATGTATTATTAAACAAGTTAATTGCCCTTTCATTAGAGAACGAATAGTGCACTCCAG 1860
Db      601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY      1861 ATCATGTGTACTTACAGAACACAGTACAAAGCCCGTCTAGCAGGGAAGCTGGCCGC 1914
Db      621 IleMetValLeuArgGlnGlnTyTrpGlnSerProSerArgGluAlaGlyArg 638

RESULT 7
QY      Q8Q6Y9_9GAMR PRELIMINARY; PRT; 641 AA.
ID      Q8Q6Y9_9GAMR PRELIMINARY; PRT; 641 AA.
AC      Q8Q6Y9;
DT      01-JUN-2002 (TrEMBLrel. 21, Created)
DT      01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Envelope glycoprotein.
GN      Name=env;
OS      Porcine endogenous retrovirus.
OC      Viruses; Retrovirdae; Retroviridae; Gammaretrovirus;
OC      1-Mammalian type C virus group.
OX      NCBI_TaxID=61673;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=21851116; PubMed=11861871;
RX      DOI=10.1128/JVI.76.6.3045-3048.2002;
RA      Oldmixon B.A., Wood J.C., Ericsson T.A., Wilson C.A.,
RA      White-Scharf M.E., Andersson G., Greenstein J.L., Schuurman H.J.,
RA      Patience C.;
RT      "Porcine endogenous retrovirus transmission characteristics of an
RT      inbred herd of miniature swine.";
RL      J. Virol. 76:3045-3048(2002).
DR      EMBL; AF417229; AAL87624.1; -; Genomic_DNA.
DR      HSSP; P03385; 1WOF.
DR      SMR; Q8Q6Y9; 491-543.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019031; C:viral envelope; IEA.
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR002050; Env polyprotein.
DR      InterPro; IPR008981; FvMvRecept-bind.
DR      Pfam; PF04229; ILV_coat; 1.
KW      Envelope protein.
SQ      SEQUENCE 641 AA; 71388 MW; 9CD82521A814923D CRC64;

Alignment Scores:
Pred. No.:      5,07e-258      Length:      641
Score:      3293.50      Matches:      614
Percent Similarity:      97.3%      Conservative:      10
Best Local Similarity:      95.8%      Mismatches:      14
Query Match:      95.0%      Indels:      3
DB:      2      Gaps:      1

US-10-723-552-3_COPY_5620_7533 (1-1914) x Q8Q6Y9_9GAMR (1-641)
QY      1 ATGCATCCACCTTAAACCGCGCCACTCCCGATTCCGGGTGGAAAGCCGAAAGACTG 60
Db      1 MethHisProThrLeuSerArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY      61 AAAATCCCTTAAAGTTTCGCTCCATCGGTGGTTCCTTACTCTCTCAATACCTCTCAG 120
Db      21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY      121 ACTAATGTGTATGCGCATAGGACAGCTGAACCTCCATAAACCTTATCTCTCACCTGG 180
Db      41 VallanGlyLysArgLeuValAsnSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY      181 TTAATTACTGACTCCGGCACAGGTATTAAATATCAACAACTCAAGGGAGAGCTCTTTA 240
Db      61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
```


Alignment Scores:

Pred. No.: 1,15e-246 Length: 653
Score: 3153.50 Matches: 594
Percent Similarity: 93.6% Conservative: 17
Best Local Similarity: 91.0% Mismatches: 27
Query Match: 90.9% Indels: 15
DB: 2 Gaps: 3

US-10-723-552-3_COPY_5620_7533 (1-1914) x Q50Q07_9GNR (1-653)

```
QY 1 ATGCATCCAGCTTAAACGGCGCCACCTCCGATTCGGGTGGAAAGCCGAAAGACTG 60
DB 1 MethHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyProLysArgLeu 20
QY 61 AAATCCCTTAAGCTTCGCCTCCATCGGTGGTTCCTTACTCTGTCAATTAACCTCTCAG 120
DB 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATGTATGCGCATAGGAGACAGCTGAACCTCCCATAAACCTTATCTCTCACCTGG 180
DB 41 ValAsnGlyLysArgLeuValAsnSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGATCCGGCACAGGTATTAAATATCAACAACTCAAGGGAGGCTCTTTA 240
DB 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTATTCCTAGTCTG----- 294
DB 81 GlyThrTrpTrpProGluLeuTrpValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 295 ---ACCTCACCCAGATATCTCCATGCTCACGGATTTTATGTTGCTCCAGGACCA 351
DB 101 GlnAlaThrProProAspValLeuArgAlaTrpGlyPheTrpValCysProGlyProPro 120
QY 352 AATAATGAAACATTCGGGAAATCCAGAGATTTCTTTGTAAACAATGGAACCTGTGTA 411
DB 121 AsnAsnGluGluTrpCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 412 ACCTTAATGATGGATATTGGAATGGCAACCTCTCAGCAGGATAGGTTAAGTTTCT 471
DB 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTrpSer 160
QY 472 TATGTCACACCATATACAGCTCTCGACAATTTAATTAC-----CTGACC 516
DB 161 PheValAsnAsnProThrSerTrpAsnGlnPheAsnTrpGlyHisGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGAACTGGAAGCCCCCAAGTGTCTCTCTTCAGAC 555
DB 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 556 CTAGATTACCTMAAATAAGTTTCACTGAGAAGGAAACAAAGAAATAATCTTAAATGG 615
DB 201 LeuAspTrpLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleLeuLysTrp 220
QY 616 GTAAATGGTATGCTCTGGGAATGCTATATATGAGGCTCGGTGTAACAACAACAGGCTCC 675
DB 221 ValAsnGlyMetSerTrpGlyMetValTrpTrpGlyLysGlySerGlyLysGlnProGlySer 240
QY 676 ATTCTAACTATTTCGCTCAAAATAAACACGAGCTGGAGCCTCCAAATGGCTATAGACCAAT 735
DB 241 IleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMetAlaIleGlyProAsn 260
QY 736 ACGGTCTTGACGGGTCAAGACCCCAACCCAGGACCAAGGACCAATCTCTTAACATAACT 795
DB 261 ThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyProSerSerAsnIleThr 280
QY 796 TCTGATCAGACCCCACTGAGTCTAACAGCAGCTAAATAATGGGGCAAACTTTTATAGC 855
DB 281 SerGlySerAspProThrGluSerAsnSerThrThrLysMetGlyAlaLysLeuPheSer 300
QY 856 CTCATCCAGGGAGCTTTTCAAGCTCTTAATCTCCAGCTCCAGAGGCTACCTCTTTGT 915
DB 301 LeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThrSerSerCys 320
```

RESULT 9

Q5J1S8_SUSBA
ID Q5J1S8_SUSBA PRELIMINARY; PRT; 642 AA.
AC Q5J1S8;

```
QY 916 TGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAAGAGGGAATTC 975
DB 321 TrpLeuCysLeuAlaLeuGlyProProTrpTrpGlyGlyMetAlaArgArgGlyLysPhe 340
QY 976 AATGTGCAAAAGAACATAGAGACCAATGCAATGGGATCCCAAAATAAGCTTACCCCTT 1035
DB 341 AsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLysLeuThrLeu 360
QY 1036 ACTGAGGTTTCTGAAAAGGACCTGCTATAGGAAGGTTCCTCCATCCCAACACACTT 1095
DB 361 ThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu 380
QY 1096 TGTAAACCACTGAAGCCTTTAATAACAACTCTCAGAGTCAATATCTGGTACCTGGTTAT 1155
DB 381 CysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTrpLeuValProGlyTrp 400
QY 1156 GACAGGTGGTGGGCATGTAATACCTGATTAACCCCTTGTGTTCACCTTGGTTTTTAAC 1215
DB 401 AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn 420
QY 1216 CAAACTAAAGATTTTTCGATTAATGCTACAGAAATCATCGACAAAGAGAGAACCCATATCT 1275
DB 421 GlnThrLysAspPheCysIleMetValGlnIleValProArgValTrpTrpProGlu 440
QY 1276 AAAGCAATCCTTGTGATGATGATGACTACAGAAATCATCGACAAAGAGAGAACCCATATCT 1335
DB 441 LysAlaIleLeuAspGluTrpAspTrpArgAsnHisArgGlnLysArgGluProIleSer 460
QY 1336 CTGACACTTGTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTAGGAAACAGGACAGCT 1395
DB 461 LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThrGlyThrAla 480
QY 1396 GCCCTGCTCAGGACCAACAGCAGCTAGAAACAGAGACTTAGTAACCTACATCGAATTGTA 1455
DB 481 AlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHisArgIleVal 500
QY 1456 ACAGAGATCTCCAAGCCCTAGAAAATCTCTAGTAACCTGGAGGAATCCTTAACCTCC 1515
DB 501 ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluLysLeuThrSer 520
QY 1516 TTATCTGAAGTAGTCTCTACAGAATAGAGGGTTAGATTTATTATTCTTAAAGAGAGA 1575
DB 521 LeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeuPheLeuLysGluGly 540
QY 1576 GGATTAATGCTAGCTTGAACGAGGAATGCTGTTTTTATGTTGATCTCAGGCGCCATC 1635
DB 541 GlyLeuCysValAlaLeuLysGluCysCysPheTrpValAspHisSerGlyAlaIle 560
QY 1636 AGAGACTCCATGAACAAGCTTAGAGAAAGTTGGAGAAGCCTCGAAGGGAAGAAAGAACT 1695
DB 561 ArgAspSerMetSerLysLeuArgGluArgLeuGluLysArgArgArgGluLysGluThr 580
QY 1696 ACTCAAGGGTGGTTTGGGGATGGTTCAACAGGTCTCTTTGTTGGCTTACCTACTTTCT 1755
DB 581 ThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeuAlaThrLeuLeuSer 600
QY 1756 GCTTTAAACAGACCCCTTAATAGTCTCTCTCTGTTACTCACAGTTGGGCCCATGTTATT 1815
DB 601 AlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrValGlyProCysIleIle 620
QY 1816 AACAGTTAAATTCCTTCATTAGAGAACGAATAAGTCAGTCAGTCACATCATGTACTTAGA 1875
DB 621 AsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMetValLeuArg 640
QY 1876 CAACAGTACCAGCCCGCTCTAGCAGGGAAGCTGGCCGC 1914
DB 641 GlnGlnTrpGlnSerProSerSerArgGluAlaGlyArg 653
```



```

Db      540  sPheTyrValAspHisSerGlyAlaIleAArgAspSerMetSerLysLeuArgGluArgLe 560
QY      1668  GGAGAGCGTGAAGGGAAGAACTACTCAAGCGTGGTTGAGGGATGTTCAACAG 1727
Db      560  uGluArgArgArgGluArgGluAaAspGlnGlyTrpPheGluGlyTrpPheAsnAr 580
QY      1728  GTCTCTTTGGTGGCTACCTACTTTCGTCTTTAACAGACCCCTTAATAGTCTCTCCT 1787
Db      580  gSerProTrpMetAlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeu 600
QY      1788  GTTACTCAGTGGGCCATGTATTATTAAACAAGTTAATTCCTTCAATAGAGAACGAAT 1847
Db      600  uLeuLeuThrValGlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgI 620
QY      1848  AAGTGCAGTCAGATCATGTGTTAGACACAGTACCAGACCCGCTTACGAGGGAAGC 1907
Db      620  eSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAl 640
QY      1908  TGGCGCG 1914
Db      640  aGlyArg 642

RESULT 10
OQ0620_9GAMR PRELIMINARY; PRT; 653 AA.
AC OQ0620_
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;
OS Porcine endogenous retrovirus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=21851116; PubMed=11861871;
RX DOI=10.1128/JVI.76.6.3045-3048.2002;
RA Oldmixon B.A., Wood J.C., Ericsson T.A., Wilson C.A.,
RA White-Schaff M.E., Andersson G., Greenstein J.L., Schuurman H.J.,
RA Patience C.;
RT "Porcine endogenous retrovirus transmission characteristics of an
RT inbred herd of miniature swine.";
RL J. Virol. 76:3045-3048(2002).
DR EMBL; AF417228; AAL87623.1; -; Genomic_DNA.
DR HSSP; P03385; IMOF.
DR SMR; OQ0620; 503-555.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env polyprotein.
DR InterPro; IPR008981; FMuLVrecept-bind.
DR Pfam; PF00429; TLV_coat; 1.
KW Envelope protein.
SQ SEQUENCE 653 AA; 73113 MW; 4A5959F988B9B3C CRC64;

Alignment Scores:
Pred. No.: 3.53e-246 Length: 653
Score: 3147.50 Matches: 593
Percent Similarity: 93.6% Conservative: 18
Best Local Similarity: 90.8% Mismatches: 27
Query Match: 90.8% Indels: 15
DB: 2 Gaps: 3

US-10-723-552-3_COPY_5620_7533 (1-1914) x OQ0620_9GAMR (1-653)
QY 1 ATGCATCCAGCTTAAACCGCGCCACCTCCGATTCCGGGTGGGAAGCCGAAGACTG 60
Db 1 MetHisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCATCGGTGGTTCCTTACTCTGTCAATAACCTCTCAG 120

```

```

Db      21  LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY      121  ACTAATGGTATGCCATAGAGACAGCCTGAACCTCCATAAACCTTATCTCTCACCTGG 180
Db      41  ValAsnGlyLysArgLeuValAsnSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY      181  TTAATTAATCTGACTCCGCGACAGAGTATTAAATATCAACACTCAAGGGAGGCTCCTTTA 240
Db      61  LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY      241  GGAACCTGGTGGCTGATATACGTTTGGCTTCAGATCAGTATTTCCTAGTCTG----- 294
Db      81  GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY      295  ---ACCTCACCCCGAGATATCTCATGCTCACCGATTTTATGTTGGCCAGGACCA 351
Db      101  GlnAlaThrProProAspValLeuArgAlaTrpGlyPheTyrValCysProGlyProPro 120
QY      352  AATAATGGAACCAATTCGGAAATCCAGAGATTTCTTTGTAAACAATGGAACCTGTGTA 411
Db      121  AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY      412  ACCTTAATGATGATATTTGAAATGGCAACCTCTCAGCAGATAGGTTAAGTTTCT 471
Db      141  ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY      472  TATGTCAACACCTATACAGCTCTGGCAATTTTAATTAC-----CTGACC 516
Db      161  PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY      517  TGG-----ATTAGAACTGGAAGCCCCAAGTCTCTCTTCAGAC 555
Db      181  TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY      556  CTAGATTACCTAAATAAAGTTTCACTGAGAAAGAAACAAAGAAATATCTCTAAATGG 615
Db      201  LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleLeuLysTrp 220
QY      616  GTAATGGTATGCTCTGGGNAATGGTATATTATATGAGGCTCGGTAAACACACAGGCTCC 675
Db      221  ValHisGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGlyLysGlnProGlySer 240
QY      676  ATTCTAACTATTTCGCTCAAAATAAAACAGCTGAGCTCCAAATGGCTTATAGGACCAAT 735
Db      241  IleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMetAlaIleGlyProAsn 260
QY      736  ACGGTCTTGAAGGCTCAAAAGACCCCAACCCCAAGACAGACCATCTCTTAACATAACT 795
Db      261  ThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyThrSerSerAsnIleThr 280
QY      796  TCTGATCAGACCCCACTGAGTCTAACAGCAGCTAAATGGGGGCGGCGGCGGCGGCGG 855
Db      281  SerGlySerAspProThrGluSerAsnSerThrThrLysMetGlyAlaLysLeuPheSer 300
QY      856  CTATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAGGCTACCTCTCTTGT 915
Db      301  LeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThrSerSerCys 320
QY      916  TGGCTATGCTTGGTTCGGGCCCACTTACTATGAAGGAATGGCTAGAGGGGGAATTC 975
Db      321  TrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAlaArgArgGlyLysPhe 340
QY      976  AATGTGCAAAAGAACATAGACCAATGCATGGGGATCCCAAAATAAGCTTACCCCTT 1035
Db      341  AsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLysLeuThrLeu 360
QY      1036  ACTGAGGTTTCTGAAAGGACCTGCTATAGAAAGGTTTCCCATCCCAACCAACCTT 1095
Db      361  ThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu 380
QY      1096  TGTAAACCACTGAAGCCTTTAATCAAAACCTCTCAGAGTCAATATCTGGTACCTGGTTAT 1155
Db      381  CysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyrLeuValProGlyTyr 400

```

QY 1156 GACAGGTGGTGGGCGATGTAATACCTGGATTAACCCCTTGTGTTCACCTGGTTTAAAC 1215
DB 401 AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn 420
QY 1216 CAAACTAAAGATTTTGCATTATGCTCCAAATGTTCCCGAGTGTTACTATCCCGAA 1275
DB 421 GlnThrLysAspPheCysIleMetValGlnIleValProArgValTyTrpProGlu 440
QY 1276 AAAGCAATCCTTGATGATATGACTACAGAAATCATCGACAAAAGAGAACCCATATCT 1335
DB 441 LysAlaIleLeuAspGluTyAspTyArgAsnHisArgGlnLysArgGluProIleSer 460
QY 1336 CTGACACTGCTGTGATGCTCGACTTGGAGTGGCAGAGGTGTAGGAACAGGACAGCT 1395
DB 461 LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThrGlyThrAla 480
QY 1396 GCCCTGGTCACGGACACACAGCAGCTAGAACAGACTTAGTACCTACATCGAATTGTA 1455
DB 481 AlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHisArgIleVal 500
QY 1456 ACAGAAGATCTCCAAAGCCTAGAAAATCTGTCACTAACCTGGAGGAATCCCTAACCTCC 1515
DB 501 ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSerLeuThrSer 520
QY 1516 TTATCTGAAGTAGTCTTACAGAAATAGAGAGGTGTAGATTTATTATTTCTAAAGAAAGGA 1575
DB 521 LeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeuPheLeuLysGluGly 540
QY 1576 GGATTATGTAGCTTGAAGGAGAAATCTGTTTATGTTGATCATTCAGGGGCCATC 1635
DB 541 GlyLeuCysValAlaLeuLysGluGluCysCysPheTyValAspHisSerGlyAlaIle 560
QY 1636 AGACACTCCATGAACAGCTTAGAGAAAGGTGTGAGAGGCTGCAAGGAGAAAGGAACT 1695
DB 561 ArgAspSerMetSerLysLeuArgGluArgLeuGluLysArgGluGluLysGluThr 580
QY 1696 ACTCAAGGTGGTTGAGGGATGTTCAACAGGCTCTCTTTGGTGGCTACCTACTTTCT 1755
DB 581 ThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeuAlaThrLeuLeuSer 600
QY 1756 GCCTTACAGGACCTTAATAGTCTCTCTCTGTACTACAGTTGGCCATGATTATT 1815
DB 601 AlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrValGlyProCysIleIle 620
QY 1816 AACAGTTAATGTCTTCATTAGAGAACGAATAGTCAGTCAGATCAGTGGTACTTAGA 1875
DB 621 AsnLysLeuIleAlaPheIleArgGluArgLysSerAlaValGlnIleMetValLeuArg 640
QY 1876 CAACAGTACCAAGCCGCTCTAGCAGGGAAGCTGGCCGC 1914
DB 641 GlnGlnTyGlnSerProSerArgGluAlaGlyArg 653

RESULT 11

Q806Z1_9GAMR
ID Q806Z1_9GAMR PRELIMINARY; PRT; 653 AA.
AC Q806Z1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;

OS Porcine endogenous retrovirus.
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=11851116; PubMed=11861871;
RX DOI=10.1128/JVI.76.6.3045-3048.2002;
RA Oldmixon B.A., Wood J.C., Ericsson T.A., Wilson C.A.,
RA White-Scharf M.E., Andersson G., Greenstein J.L., Schuurman H.J.,
RA Patience C.;

RT "Porcine endogenous retrovirus transmission characteristics of an
RL inbred herd of miniature swine.";
J. Virol. 76:3045-3048(2002).
DR EMBL; AF417227; AAL87622.1; -; Genomic_DNA.
DR HSSP; P01385; INOF.
DR SMR; Q8Q6Z1; 503-555.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env_polyprotein.
DR InterPro; IPR008981; FMuLVrecept-bind.
DR Pfam; PF00429; TLV_coat; 1.
KW Envelope protein.
SQ SEQUENCE 653 AA; 73056 MW; 9E9D00137DDBA244 CRC64;

Alignment Scores:

Pred. No.: 3,32e-245 Length: 653
Score: 3135.50 Matches: 591
Percent Similarity: 93.1% Conservative: 17
Best Local Similarity: 90.5% Mismatches: 30
Query Match: 90.4% Indels: 15
DB: 2 Gaps: 3

US-10-723-552-3_COPY_5620_7533 (1-1914) x Q806Z1_9GAMR (1-653)

QY 1 ATGCATCCACGTTAAACCGCGCCACCTCCCGATTCGGGTGGAAAGCCGAAAGACTG 60
DB 1 MethHisProThrLeuSerArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAGTTCGCTCATCGCTGGTTCCTTACTCTGTCAATAACTCTCAG 120
DB 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATGGTATGCGATAGGAGCAGCTGACCTCCATACCACTCAAGGGAGGCTCTTTA 180
DB 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACCTCCGCGCACAGGTATTAAATATCAACACACTCAAGGGAGGCTCTTTA 240
DB 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATAGTTTGCTTCAGATCAGTATTCTAGTCTG----- 294
DB 81 GlyThrTrpTrpProGluLeuTyValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 295 ---ACCTCACCCCGCATATCCCTCCATGCTCAGGATTTTATGTTGCCAGGACCA 351
DB 101 GlnAlaThrProProAspValLeuArgAlaTyGlyPheTyValCysProGlyProPro 120
QY 352 AATAATGGAAACAVTTCGGAAATCCGAGATTTCTTTTGTAAACATGGAACTGTGTA 411
DB 121 AsnAsnGluGluTyCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 412 ACCTCTAATGATGATTTGGAATGCGCAACTCTTCAGCAGGATAGGGTAAGTTTCT 471
DB 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTySer 160
QY 472 TATGTCACACCTATACCGCTCTGGACATTTAATTAC-----CTGACC 516
DB 161 PheValAsnAsnProThrSerTyAsnGlnPheAsnTyGlyHisGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGAACTGGAAGCCCAAGTCTCTCTCTCTAGAC 555
DB 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 556 CTAGATTACCTAAAAATAAGTTTTCATCTGAGAAAGGAAAAACAAGAAATATCTTAAATGG 615
DB 201 LeuAspTyLeuLysIleSerPheThrGluLysGlyGlnGlnAsnIleLeuLysTrp 220
QY 616 GTAAATGGTATGTTGGGGAATGGTATATTATGGAGGCTCGGTTAAACACCAAGGCTCC 675
DB 221 ValAsnGlyMetSerTrpGlyMetValTyTrpGlySerGlyLysGlnProGlySer 240

```
Qy 676 ATTCTAACTATTGCGCTCAAAATAAACAGCTGGAGCTCCAAATGGCTATAGGACCAAT 735
Db |||||
Qy 241 IleuThrIleArgLeuGlyAsnGlnLeuGluProPheAlaIleGlyProAsn 260
Db |||||
Qy 736 ACGGTCTTGACGGGTCAAAGACCCCAACCAAGGACCAGGACCATCTCTAAACATACT 795
Db |||||
Qy 261 ThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyProSerSerAsnIleThr 280
Db |||||
Qy 796 TCTGATCAGACCCCACTGAGCTTAAACAGCAGCTAAATGGGGGCAAACTTTTAGC 855
Db |||||
Qy 281 SerGlySerAspProThrGluSerAsnSerThrThrLysMetGlyAlaLysLeuPheSer 300
Db |||||
Qy 856 CTGATCCAGGAGCTTTTCAAGCTCTTAACTCCAGACTCCAGAGCTCCAGAGCTTCTTGT 915
Db |||||
Qy 301 LeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThrSerSerCys 320
Db |||||
Qy 916 TGGCTATGCTTAGCTTCGGGCCCACTTACTATGAAGGAATGGCTAGAGAGGGAATTC 975
Db |||||
Qy 321 TrpLeuCysLeuAlaLeuGlyProProThrTyrrGluGlyMetAlaAArgGlyLysPhe 340
Db |||||
Qy 976 AATGTGACAAAGAACATAGACCAATGCACATGGGGATCCCAAAATAAGCTTACCTT 1035
Db |||||
Qy 341 AsnValThrLysGlyHisArgAspProCysThrTrpGlySerGlnAsnLysLeuThrLeu 360
Db |||||
Qy 1036 ACTGAGGTTTCTGGAARAGGCACCTGCATAGAGAGGTTCCCCATCCCAACACACTT 1095
Db |||||
Qy 361 ThrGluValPheGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu 380
Db |||||
Qy 1096 TGTAAACCACTGAAGCCCTTTAATCAAACTCTGAGAGTCAATATCTGTGTAACCTGGTTAT 1155
Db |||||
Qy 381 CysAsnHisThrGluAlaPheAsnArgThrSerGluSerGlnTyrrLeuValProGlyTyrr 400
Db |||||
Qy 1156 GACAGGTGGTGGGCATGTAATACCTGAGTAACCCCTTGTTGTTCCACTTGGTTTTAAC 1215
Db |||||
Qy 401 AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn 420
Db |||||
Qy 1216 CAACCTAAGATTTTGCATTATGTCCTCAAAATGTTCCCGAGTCTATTACTATCCCGAA 1275
Db |||||
Qy 421 GlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrrTyrrProGlu 440
Db |||||
Qy 1276 AAAGCAATCCTTGATGATATGACTACAGAAATCATCGACAAAAGAGAACCCATATCT 1335
Db |||||
Qy 441 LysAlaIleLeuAspGluTyrrAspTyrrArgAsnHisArgGlnLysArgGluProIleSer 460
Db |||||
Qy 1336 CTGACACTTGCTGTGATCTCGACTCGAGTGGCAGCAGGTGTAGGAACAGGACAGCT 1395
Db |||||
Qy 461 LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleValGlyThrGlyThrAla 480
Db |||||
Qy 1396 GCCCTGGTCACGGGACACAGCAGCTAGAAACAGGACTTAGTAACTACATCGAATTGTA 1455
Db |||||
Qy 481 AlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHisArgIleVal 500
Db |||||
Qy 1456 ACAGAAATCTCCAGCCCTAGAAAATCTGTCTAGTAACTCGAGGAATCCCTAACCTCC 1515
Db |||||
Qy 501 ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluSerLeuThrSer 520
Db |||||
Qy 1516 TTAATCTGAAGTAGTCTCAGAAATAGAGAGGTTAGATTTATTATTCTTAAAGAGGA 1575
Db |||||
Qy 521 LeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeuPheLeuLysGluGly 540
Db |||||
Qy 1576 GGATTTATGTGTAGCTTCAAGGAGGATGCTGTTTTTTATGTTGATCATTCAGGGGCCATC 1635
Db |||||
Qy 541 GlyLeuCysValAlaLeuLysGluGluCysCysPheTyrrValAspHisSerGlyAlaIle 560
Db |||||
Qy 1636 AGAGACTCCATGAACACTTAGAGAAAGGTTGGAGAGCGCTCGAAGGGGAAAAGAACT 1695
Db |||||
Qy 561 ArgAspSerMetSerLysLeuArgGluArgLeuGluLysArgArgGluGluGluThr 580
Db |||||
Qy 1696 ACTCAAGGGTGGTTGAGGATGGTTCAACAGGTCTCTTTGGTGGCTTACCCCTACTTTCT 1755
Db |||||
Qy 581 ThrGlnGlyTyrrPheGluGlyTyrrPheAsnArgSerProTrpLeuAlaThrLeuLeuSer 600
Db |||||
Qy 1756 GCTTTAAACAGGACCCCTTAATAGTCTCTCTCTGTTTACTCACAGTGGGGCCATGATTATT 1815
Db |||||
```

```
Db 601 AlalaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrValGlyProCysIleIle 620
Qy 1816 AACAGTAATATGCTTCATTAGAGAACGAATAAGTCAGTCGATCATCGTACTTAGA 1875
Db |||||
Qy 621 AsnLysLeuIleAlaPheIleArgGluArgIleAlaValGlnIleMetValLeuArg 640
Db |||||
Qy 1876 CACAGTACCAAGCCCGTCTAGCAGGGAAGCTGGCCGC 1914
Db |||||
Qy 641 GlnGlnTyrrGlnSerProSerSerArgGluAlaGlyArg 653
Db |||||
RESULT 12
Q8Q6Y8_9GAMR
ID Q8Q6Y8_9GAMR PRELIMINARY; PRT; 659 AA.
AC Q8Q6Y8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Envelope glycoprotein.
GN Name=env;
OS Porcine endogenous retrovirus.
OC Viruses; Retrovirdae; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21851116; PubMed=11861871;
RX DOI=10.1128/JVI.76.6.3045-3048.2002;
RA Oldmixon B.A., Wood J.C., Ericsson T.A., Wilson C.A.,
RA White-Scharf M.E., Andersson G., Greenstein J.L., Schuurman H.J.,
RA Patience C.;
RA "Porcine endogenous retrovirus transmission characteristics of an
RT indred herd of miniature swine.";
RL J. Virol. 76:3045-3048(2002).
DR EMBL; AF417230; AAL87625.1; -; Genomic_DNA.
DR HSSP; P03385; 1MOF.
DR SMR; Q8Q6Y8; 509-561.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env_Polyprotein.
DR InterPro; IPR008981; FmULVrecept-bind.
DR Pfam; PF00429; TLV_coat; 1.
KW Envelope protein.
SQ SEQUENCE 659 AA; 73902 MW; 1D0C1D84466B52F CRC64;
Alignment Scores:
Pred. No.: 1.02e-235 Length: 659
Score: 3018.50 Matches: 575
Percent Similarity: 90.4% Conservative: 21
Best Local Similarity: 87.3% Mismatches: 42
Query Match: 87.0% Indels: 21
DB: 2 Gaps: 6
US-10-723-552-3_copy_5620_7533 (1-1914) x Q8Q6Y8_9GAMR (1-659)
Qy 1 ATGCATCCACCTTAACCCGCGCACCTCCGATTCGGGTGGAAGCCGAAGACTG 60
Db |||||
Qy 1 MethisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu 20
Db |||||
Qy 61 AAAATCCCTTAACTTCGCCTCCATCGCGTTCCTTACTCTGTCAATAACCTCTCAG 120
Db |||||
Qy 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
Db |||||
Qy 121 ACTAATGATGTCGATAGGACAGCGCTGAACCTCCATAAACCTTATCTCTCACCTGG 180
Db |||||
Qy 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
Db |||||
Qy 181 TTAATTAATGATCGCGCACAGGTATTAAATATCAACAACACTCAAGGGAGGCTCTTTA 240
Db |||||
Qy 61 LeuIleThrAspSerGlyThrGlyLeuAsnIleAsnSerThrGlnGlyAlaProLeu 80
Db |||||
Qy 241 GGAACCTGGTGGCCTGATCTATACGTTTGGCTCAGATCAGTTATTCTTAGTCTG----- 294
Db |||||
```


Db	441	ValTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGln	460
Qy	1318	AAGAGAAACCCATATCTCTGACACTTGTGTGATGTCGCGACTTGGAGTGGCAGCAGGT	1377
Db	461	LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly	480
Qy	1378	GTAGGAACAGGAACAGCTGCCCTGGTCACGGGACACACAGAGCTAGAAACAGAGCTTAGT	1437
Db	481	ValGlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSer	500
Qy	1438	AACCTACATCGAATTGTAAACAGAAGACTCTCCAAGCCCTAGAAAAATCTGTCAGTAACCTG	1497
Db	501	AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu	520
Qy	1498	GAGGAATCCCTAACCTCTTATCTGAAGTAGTCCTACAGAAATAGAAGAGGTTAGATTTA	1557
Db	521	GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu	540
Qy	1558	TTATTTCTAAAGAGAGGATTATGTAGTCCTTGAAGGAGGAATGCTGTTTTTATGTC	1617
Db	541	LeuPheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal	560
Qy	1618	GATCATTTCAGGGGCCATCAGAGACTCCATCAAGCAAGCTTAGAGAAAAGTTGGAGAGAGCGT	1677
Db	561	AspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluLysArg	580
Qy	1678	CGAAGGAAAGAAACTACTCAAGGGTGTTTGAAGGATGGTTCAACAGAGTCTCTTTGG	1737
Db	581	ArgArgGluLysGluThrThrGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyr	600
Qy	1738	TTGGCTACCTACTTCTCTGTTTAAACAGGACCTTAATAGTCTCTCTCTGTTTACTCACA	1797
Db	601	LeuAlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThr	620
Qy	1798	GTTGGCCCATGTTATTAAACAGTTAATGCTTCATTAGAGAAAGTAAGTGCAGTC	1857
Db	621	ValGlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaVal	640
Qy	1858	CAGATCATGTACTTAGACAAACAGTACAAAGCCCTGTAGCAGGGAAGCTGGCCCG	1914
Db	641	GlnIleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg	659
RESULT 13			
Q8Q6Y6_9GAMR PRELIMINARY; PRT; 659 AA.			
AC	Q8Q6Y6		
AD	Q8Q6Y6		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	Envelope glycoprotein.		
GN	Name=env;		
OS	Porcine endogenous retrovirus.		
OC	Viruses; Retroid viruses; Retroviridae; Gammaretrovirus;		
OC	1-Mammalian type C virus group.		
OX	NCBI_TaxID=61673;		
ON	[1]		
RP	NECLTOTIDE SEQUENCE.		
RP	MEDLINE=21851116; PubMed=11861871;		
RX	DOI=10.1128/JVI.76.6.3045-3048.2002;		
RA	Oldmixon B.A., Wood J.C., Ericsson T.A., Wilson C.A.,		
RA	White-Scharf M.E., Andersson G., Greenstein J.L., Schuurman H.J.,		
RA	Patience C.;		
RT	"Porcine endogenous retrovirus transmission characteristics of an		
RT	inbred herd of miniature swine.";		
RL	J. Virol. 76:3045-3048(2002).		
RL	EMBL; AF417232; AAL87627.1; -; Genomic_DNA.		
DR	HSSP; P03385; IMOF.		
DR	SMR; Q8Q6Y6; 509-561.		
DR	GO; GO:0019028; C:viral capsid; IEA.		
DR	GO; GO:0019031; C:viral envelope; IEA.		
DR	GO; GO:0005198; F:structural molecule activity; IEA.		
DR	InterPro; IPR002050; Env_polyprotein.		
DR	InterPro; IPR008981; FmULVrecept-bind.		

RESULT 14
 ID Q8Q6Y7_9GAMR PRELIMINARY; PRT; 659 AA.
 AC Q8Q6Y7;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Envelope glycoprotein.
 GN Name=env;
 OS Porcine endogenous retrovirus.
 OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus;
 OC 1-Mammalian type C virus group.
 OX NCBI_TaxID=61673;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21851116; PubMed=11861871;
 RX DOI=10.1128/JVI.76.6.3045-3048.2002;
 RA Oldmixon B.A., Wood J.C., Ericsson T.A., Wilson C.A.,
 RA White-Scharf M.E., Andersson G., Greenstein J.L., Schuurman H.J.,
 RA Patience C.;
 RT "Porcine endogenous retrovirus transmission characteristics of an
 RT inbred herd of miniature swine.";
 RL J. Virol. 76:3045-3048(2002).
 DR EMBL; AF417231; AAL87626.1; -; Genomic_DNA.
 DR HSP; P03385; IMOF.
 DR SMR; Q8Q6Y7; 509-561.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR002050; Env_polyprotein.
 DR InterPro; IPR008981; FmULVrecept-bind.
 DR Pfam; PF00429; TIV_coat; 1.
 KW Envelope protein.
 SQ SEQUENCE 659 AA; 73867 MW; 7B6A58D11541CA06 CRC64;

Alignment Scores:
 Pred. No.: 7,51e-234 Length: 659
 Score: 2995.50 Matches: 569
 Percent Similarity: 90.1% Conservative: 25
 Best Local Similarity: 86.3% Mismatches: 44
 Query Match: 86.4% Indels: 21
 DB: 2 Gaps: 6

US-10-723-552-3_COPY_5620_7533 (1-1914) x Q8Q6Y7_9GAMR (1-659)

QY 1 ATGCATCCACGTTAAACCGCGCACCTCCCGATTCCGGGTGAAAGCGAAAGACTG 60
 DB 1 MethisProThrLeuSerArgArgHisLeuProLeuArgGlyLysProLysArgLeu 20
 QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTTCCTTACTCTGTCAATAACCTCTCAG 120
 DB 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
 QY 121 ACTAATGGTATGCCATAGGACAGCAGCTGACCTCCCATAAACCTTATCTCTCACCTGG 180
 DB 41 IleAsnGlyLysArgLeuValSerProAsnSerHisLysProLeuSerLeuThrTrp 60
 QY 181 TTAATTAAGTCCCGCACAGGTATTATATCAACACACTCAAGGGAGGCTCCCTTA 240
 DB 61 LeuLeuThrAspSerGlyThrGlyIleAsnLeuSerThrGlnGlyGluAlaProLeu 80
 QY 241 GGAACCTGGTGGCTGATCTATAGTTTGCTTCAGATCAGTTATCTCTAGTCTG----- 294
 DB 81 GlyThrTrpTrpProGluLeuLeuValCysLeuArgSerValIleProGlyLeuAsnAsp 100
 QY 295 ----ACCTCACCCAGATATCTCCATGCTCAGATTTTATGTTTGGCCAGGACCA 351
 DB 101 GlnAlaThrProAspValLeuArgAlaTrpGlyPheTrpValCysProGlyProPro 120
 QY 352 AATAATGGAACATTCGGAATCCAGAGATTTCTTTGTAACAAATGGAATGTTA 411
 DB 121 AsnAsnGluGluTrpCysGlyAsnProGlnAspPheCysLysGlnTrpSerCysVal 140

QY 412 ACCTCTAATGATGATATTGGAAATGCCAACCTCTCAGCAGGATAGGTTTCT 471
 DB 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
 QY 472 TATGTCACACCTATACAGCTCGGACAAATTTAATTAC-----CTGACC 516
 DB 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
 QY 517 TGG-----ATTGAACTGGAAGCCCAAGTCTCTCTCTTCAGAC 555
 DB 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
 QY 556 CTAGATTACCTAAATAAGTTTCACCTGAGAAAGGAAACAAAGAAAATATCTCTAAAATGG 615
 DB 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrp 220
 QY 616 GTAAATGGTATGCTTGGGNAATGTTATATATGAGGCTCGGTAAACACACGAGCTCC 675
 DB 221 ValAsnGlyMetSerTrpGlyIleValTyrGlySerGlyArgLysGlySer 240
 QY 676 ATTCTAATATTGCGCTCAAAATA---AACCAGCTGGAGCCTCCAATGGCTATAGGACCA 732
 DB 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProProValAlaIleGlyPro 260
 QY 733 AATACGGTCTTGACGGGTCAAGACCCCAACCCAA-----GGACCAAGGACCA 780
 DB 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGluArgProSerProAsnPro 280
 QY 781 TCCTCT---NACATAACTTCTGGATCAGACCCCACTAGTCTAACACGACGACTAAAATG 837
 DB 281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
 QY 838 GGGGCAAAACTTTTAGCCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCACGACTCCA 897
 DB 301 GlyAlaLysLeuPheAsnLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
 QY 898 GAGGCTACCTCTTTTGTTCGCTATGCTTAGCTTCGGGCCCACTTACTATGAAGGAATG 957
 DB 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMet 340
 QY 958 GCTAGAAGAGGGAATTCATCTGACAAAAGACATAGACACCATCATCGGGATCC 1017
 DB 341 AlaArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
 QY 1018 CAAATAAGCTTACCTTACTAGGTTCTCGAAAAGCGACCTGCATAGAAAAGTTCC 1077
 DB 361 GluAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValPro 380
 QY 1078 CCATCCCAACCAACCTTTGTAAACCACTGAAGCCTTTAATCAAACTCTGAGAGTCAA 1137
 DB 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGln 400
 QY 1138 TATCTGTACTCGTTATGACAGGTGGTGGCATGTAATTAATGATTAACCCCTTGTT 1197
 DB 401 TyrLeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 420
 QY 1198 TCCACCTTGTTTAAACCAAACTTAAAGATTTTTCATTATGTCCTCAAAATTTGTC 1257
 DB 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
 QY 1258 GTGTATTACTATCCGAAAAGCAATCTTGTATGAATATGACTACAGAAATCATCGACAA 1317
 DB 441 ValTyrTyrTrpProGluLysAlaValLeuAspGluTyrAspTyrArgAsnHisArgGln 460
 QY 1318 AAGAGAGAACCAATATCTCTGACACTTCTGTGATGCTCGGACTTGGAGTGGCGAGGTT 1377
 DB 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
 QY 1378 GTAGGAACAGGAACAGCTGCCCTGGTCACGGGACACAGCAGCTAGAAAACAGACTTAGT 1437
 DB 481 ValGlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSer 500
 QY 1438 AACCTACATCGAATTGTAAACAGAAGATCTCCAAGCCCTAGAAAAATCTGTCAGTAACCTG 1497

```

Db      501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY      1498 GAGGAATCCCTAAACCTCTTATCTGAAGTAGCTCTACAGATAGAGAAGGTTAGATTTA 1557
Db      521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY      1558 TTATTTCTAAAGAGAGGAGGATTATGTAGCCTTGAAGGAGGATGCTGTTTATGTG 1617
Db      541 LeuPheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal 560
QY      1618 GATCATTCAGGGGCCATCAGAGACTCCATCAACAGCTTAGAGAAAGGTTGGAGAAGCGT 1677
Db      561 AspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArg 580
QY      1678 CGAAGGGAAGAACTACTCAAGGTGGTTTGAGGGATGTTCAACAGTCTCTTTGG 1737
Db      581 ArgArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY      1738 TTGGCTACCTTCTTCTGCTTAAACAGCCCTTAACAGAGCCCTTAATAGTCCCTCTCTTACTCACA 1797
Db      601 LeuAlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThr 620
QY      1798 GTTGGCCATGTATTATTAAACAAGTTAATTGCCTTCATTAGAGAACGAATAAGTGCAGTC 1857
Db      621 ValGlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaVal 640
QY      1858 CAGATCATGTGTTAGTACAAACAGTACCAAGCCCGTCTAGCAGGAAGCTGGCGGC 1914
Db      641 GlnIleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 659

```

RESULT 15

```

Q6UQ77_9GAMR
ID Q6UQ77_9GAMR PRELIMINARY; PRT; 659 AA.
AC Q6UQ77;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Envelope.
GN Name=env;
OS Porcine endogenous retrovirus.
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus;
OC 1-Mammalian type C virus group.
OX NCBI_TaxID=61673;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14963152; DOI=10.1128/JVI.78.5.2502-2509.2004;
RA Scobie L., Taylor S., Wood J.C., Suling K.M., Quinn G., Meikle S.,
RA Patience C., Schuurman H.J., Onions D.E.;
RT "Absence of replication-competent human-tropic porcine endogenous
RT retroviruses in the germ line DNA of inbred miniature swine.";
RL J. Virol. 78:2502-2509(2004).
DR EMBL; AY364236; AAR16717.1; -; Genomic_DNA.
DR HSSP; P03390; 1AOL.
DR SMR; Q6UQ77; 509-561.
DR GO; GO:0039028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002050; Env polyprotein.
DR Pfam; PF00429; TUV coat; 1.
SQ SEQUENCE 659 AA; 74068 MW; 466F54D8CF93014F CRC64;

```

Alignment Scores:

```

Pred. No.: 4,876-230 Length: 659
Score: 2948.50 Matches: 562
Percent Similarity: 89.1% Conservative: 25
Best Local Similarity: 85.3% Mismatches: 51
Query Match: 85.0% Indels: 21
DB: 2 Gaps: 6

```

US-10-723-552-3_COPY_5620_7533 (1-1914) x Q6UQ77_9GAMR (1-659)

```

QY      1 ATGCATCCACGCTTAAACCGCGCCACCTCCCGATTCCGGGTGGAAAGCCGAAAGACTG 60
      ||||| :|||:|||||

```

```

Db      1 MethHisSerThrLeuSerArgArgHisPheProileArgGlyGlyLysProLysArgLeu 20
QY      61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCTTACTCTGTCAATAACCTCTCAG 120
Db      21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY      121 ACTAATGCTATGCGCATAGGACAGCGCTGAACCTCCCATAAACCTTATCTCTCACCTGG 180
Db      41 ValAsnGlyLysArgLeuValAspCysProAsnSerHisLysProLeuSerLeuThrTrp 60
QY      181 TTAATTACTGACTCCGCGCACAGGTATTAAATATCAACAACACTCAAGGAGGAGCTCTTTA 240
Db      61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY      241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTATTATCTAGTCGTG----- 294
Db      81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY      295 --ACCTCACCCCGACATATCTCCATGCTCACCAGATTTTATGTTTGGCCAGGACCACCA 351
Db      101 GlnAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
QY      352 AATAATGAAACAATTCGGAAATCCAGAGATTTCTTTCTTAAACAATGAACTGTGTA 411
Db      121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY      412 ACCTCTAATGATGATATTGGAATGGCCACCTCTCAGCAGGATAGGTAACTTTTCT 471
Db      141 ThrSerAsnAspGlyAsnTrpLysTrpValSerGlnGlnAspArgValSerTyrSer 160
QY      472 TATGTCAACCTATACCAGCTCTGGCAATTTAATTAC-----CTGACC 516
Db      161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY      517 TGG-----ATTAGAACTGGAAGCCCAAGTGTCTCTCTTCAGAC 555
Db      181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY      556 CTAGATTACCTAAATAAGTTTCACTCAGAAAGGAACAAGAAATATCTCTAAATGG 615
Db      201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrp 220
QY      616 GTAAATGCTATGCTTGGGGAATGATATTATGAGGCTCGGGTAAACAACACAGGCTCC 675
Db      221 ValAsnGlyMetSerTrpGlyIleValTyrTyrGlyGlySerGlyArgLysGlySer 240
QY      676 ATTCCTAATTCCTCCTCAAAATA---AACAGCTGGAGCTCCAAATGGCTATAGGACCA 732
Db      241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProProValAlaIleGlyPro 260
QY      733 AATAGGCTCTTGACGGGTCAAGAGACCCCAACCCCAAGGA-----CCAGGACCA 780
Db      261 AsnLysGlyLeuAlaLysGlnGlyProProIleGlnGluArgProPheProAsnPro 280
QY      781 TCCTCT---AACATAACTTCTGGATCAGACCCCACTGAGTCTAACAGCACCACCTAAATG 837
Db      281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY      838 GGGGCAAACTTTTTCAGCTCATCCAGGAGCTTTTCAAGCTCTTTAACTCCACGACTCCA 897
Db      301 GlyAlaLysLeuPheAsnLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY      898 GAGGCTACCTCTTCTTGTGGTATGCTTACGTTCCGGCCCACTTACTATGATGAGGAATG 957
Db      321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMet 340
QY      958 GCTAGAGAGAGGAATTCATATGTGACAAAGAAACATAGAGACCAATGACATCGGGATCC 1017
Db      341 AlaArgGlyGlyArgPheTyrValThrLysGluHisArgAspGlnCysAlaTrpGlySer 360
QY      1018 CAAATAAGCTTACCTTACTGAGTTTCTGGAAAAGGCACCTGCATAGGAAAGGTTCC 1077
Db      361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380

```

QY 1078 CCATCCCAACCAACCTTTGTAACCACTGAAGCCTTTAATCAAACTCTGAGAGTCAA 1137
Db |||||||
381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 1138 TATCTGCTACCTGTTTATGACAGGTGGTGGCATGTATTAATCTGGATTAACCCCTTGTTGTT 1197
Db |||||||
401 TyrLeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 1198 TCCACCTTGGTTTTTAAACCAAACTAAAGATTTTTGTCATTATGGTCCAAATTTGCCCGA 1257
Db |||||||
421 SerThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArg 440
QY 1258 GTGTATTACTATCCGAAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAA 1317
Db |||||||
441 ValTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGln 460
QY 1318 AAGAGAGAACCCATATCTCAGCAGCTGCTGATGCTCGGACTTGGAGTGGCAGCAGGT 1377
Db |||||||
461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
QY 1378 GTAGGAACAGGAACAGCTGCCCTGGTCACGGGACCACAGCAGCTAGAAACAGGACTAGT 1437
Db |||||||
481 ValGlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSer 500
QY 1438 AACCTACATCGAATTGTTAACAGAGATCTCCAAGCCCTAGAAAAATCTGTCAGTAACCTG 1497
Db |||||||
501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 1498 GAGGAATCCCTAACCTCTTATCTGAAGTAGTCTCCTACAGATAGAAAGGGTTAGATTTA 1557
Db |||||||
521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 1558 TTATTTCTAAAGAGGAGGATTATGTAGCTTTGAAGGAGGAATGCTGTTTTTATGTG 1617
Db |||||||
541 LeuPheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal 560
QY 1618 GATCATTGAGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAAGGTTGAGAGCGGT 1677
Db |||||||
561 AspHisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArg 580
QY 1678 CGAAGGGAAAGGAACCTACTCAAGGGTGGTTGAGGGATGGTTCAACAGGTCTCTTTGG 1737
Db |||||||
581 ArgArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 1738 TTGGCTACCTTACTTCTGCTTTAACAGGACCTTATAGTCTCTCTCTCTACTCACA 1797
Db |||||||
601 LeuAlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThr 620
QY 1798 GTTGGGCCATGTATTATTAAACAGTTAATTCCTTTTCAATTAGAGAACGAATAAGTGCAGTC 1857
Db |||||||
621 ValGlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaVal 640
QY 1858 CAGATCATGTGTACTTAGACAACAGTACCAGCCGCTCTAGCAGGGAGCTGGCCGC 1914
Db |||||||
641 GlnIleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 659

Search completed: February 14, 2006, 14:51:01
Job time : 448.173 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 12:52:14 ; Search time 50.9239 Seconds
(without alignments)
3302.856 Million cell updates/sec

Title: US-10-723-552-3_COPY_5620_7533

Perfect score: 3468

Sequence: 1 ATGCATCCACGTTAAACCG.....CTACGAGGAGAGCTGGCCGC 1914

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB_spool/US10723552/runat_14022006_125139_12833/app_query.fasta_1
-DB-A Geneseq -QFWT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=dits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10723552 @CGN 1.1.734 @runat_14022006_125139_12833 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A Geneseq 21:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3425	98.8	638	2 AAW32098	AAW32098 Miniature
2	3425	98.8	638	4 AAB70633	AAB70633 Porcine e
3	3425	98.8	638	4 AAB73287	AAB73287 Retroviral
4	3425	98.8	638	8 ADS73448	ADS73448 Swine ret
5	3420	98.6	638	4 AAB35114	AAB35114 PERV-A en
6	3407	98.2	638	9 ADY28028	ADY28028 Porcine e
7	3407	98.2	638	9 ADY28030	ADY28030 Porcine e
8	3407	98.2	638	9 ADY28036	ADY28036 Porcine e
9	3407	98.2	638	9 ADY28041	ADY28041 Porcine e

10	3407	98.2	638	9	ADY28038	ADY28038 Porcine e
11	3402	98.1	638	4	AAB70634	AAB70634 Porcine e
12	3402	98.1	666	8	ADS73449	ADS73449 Swine ret
13	3318	95.7	678	9	ADY28032	ADY28032 Porcine e
14	3318	95.7	678	9	ADY28034	ADY28034 Porcine e
15	3147.5	90.8	653	6	ABB82635	ABB82635 PERV enve
16	3135.5	90.4	653	6	ABB82636	ABB82636 PERV enve
17	2940.5	84.8	659	4	AAB35112	AAB35112 PERV-1-15
18	2879.5	83.0	661	9	ADY28039	ADY28039 Porcine e
19	2874.5	82.9	660	9	ADY28012	ADY28012 Porcine e
20	2871.5	82.8	660	9	ADY28004	ADY28004 Porcine e
21	2866.5	82.7	660	2	AAW85453	AAW85453 Pig endog
22	2866.5	82.7	660	2	AAW85453	AAW85453 Pig endog
23	2866.5	82.7	660	4	AAB70632	AAB70632 PERV-A en
24	2851	82.2	678	9	ADY28008	ADY28008 Porcine e
25	2849.5	82.2	660	9	ADY28014	ADY28014 Porcine e
26	2849.5	82.2	660	9	ADY28006	ADY28006 Porcine e
27	2835	81.7	678	9	ADY28010	ADY28010 Porcine e
28	2755.5	79.5	657	2	AAW32095	AAW32095 Porcine e
29	2755.5	79.5	657	8	ADS73445	ADS73445 Swine ret
30	2755.5	79.5	1069	4	AAB73284	AAB73284 Defective
31	2465	71.1	678	9	ADY28022	ADY28022 Porcine e
32	2455	70.8	678	9	ADY28020	ADY28020 Porcine e
33	2432	70.1	657	9	ADY28026	ADY28026 Porcine e
34	2432	70.1	657	9	ADY28018	ADY28018 Porcine e
35	2432	70.1	658	9	ADY28040	ADY28040 Porcine e
36	2422	69.8	657	2	AAW85452	AAW85452 Pig endog
37	2422	69.8	657	4	AAB70635	AAB70635 Porcine e
38	2422	69.8	657	9	ADY28024	ADY28024 Porcine e
39	2422	69.8	657	9	ADY28016	ADY28016 Porcine e
40	2395	69.1	657	2	AAW39274	AAW39274 Porcine e
41	2353	67.8	656	2	AAW39273	AAW39273 Porcine e
42	2353	67.8	656	4	AAB70636	AAB70636 Porcine e
43	1989	57.4	540	6	ABP58412	ABP58412 Porcine e
44	1985	57.2	540	6	ABP58415	ABP58415 Porcine e
45	1985	57.2	540	6	ABP58413	ABP58413 Porcine e

ALIGNMENTS

RESULT 1

AAW32098

ID AAW32098 standard; protein; 638 AA.

XX AC AAW32098;

XX AC AAW32098;

DT 27-AUG-2003 (revised)

DT 09-FEB-1998 (first entry)

XX DT 09-FEB-1998 (first entry)

DE Miniature swine retrovirus ENV protein.

XX DE Miniature swine retrovirus ENV protein.

KW Retrovirus; porcine; ENV protein; xenotransplantation; infectious;

KW provirus; organ transplant; donor; activated virus; PCR.

XX OS Pig endogenous retrovirus.

XX OS Pig endogenous retrovirus.

PH Key Location/Qualifiers

FT Protein 1..638

FT /label= ENV_protein

XX WO921836-A1.

XX WO921836-A1.

XX PD 19-JUN-1997.

XX PD 19-JUN-1997.

XX PF 13-DEC-1996; 96WO-US019680.

XX PF 13-DEC-1996; 96WO-US019680.

XX PR 14-DEC-1995; 95US-00572645.

XX PR 14-DEC-1995; 95US-00572645.

XX PA (GEO) GEN HOSPITAL CORP.

XX PA (GEO) GEN HOSPITAL CORP.

XX PI Fishman JA;

XX PI Fishman JA;

XX DR WPI; 1997-332804/30.

XX DR WPI; 1997-332804/30.

DR N-PSDB; AAT74884.
XX New nucleic acid from porcine retro:viruses - used for detecting viruses
PT in transplant or other tissue and for assessing risk of transmitting
PT infection to graft recipient.
XX
PS Claim 22; Fig 3; 128pp; English.

XX This is a porcine retrovirus from miniature swine containing the coding
CC region for a putative viral ENV protein. This sequence and PCR fragments
CC generated from the sequence (see AAT74812-T74892) could be used to screen
CC organs for porcine retroviruses prior to xenotransplantation.
CC Transplantation can increase the likelihood of retroviral activation if
CC intact and infectious proviruses are present. The porcine retroviral
CC sequence can be used to generate probes to determine the level (e.g. copy
CC number) of intact (i.e. potentially replicating) porcine provirus
CC sequences in a strain of xenograft transplantation donors. It can be used
CC to detect mutations, genetic lesions or viral recombinants and also to
CC determine the histological localisation of activated retrovirus. Using
CC Polymerase Chain Reaction DNA Quantitation (PQD) on blood mononuclear
CC cells, infectivity titration and susceptibility testing can be performed.
CC ultimately animal donors without intact porcine retroviral sequences or a
CC lower copy number of viral elements could be selected. (Updated on 27-AUG
CC -2003 to correct OS field.)
XX

SQ Sequence 638 AA;

Alignment Scores:

Pred. No.: 2.8e-291 Length: 638
Score: 3425.00 Matches: 638
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.8% Indels: 0
DB: 2 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x AAW32098 (1-638)

QY 1 ATGCATCCAGCTTAACCGGCGCCACCTCCGATTCGGGTGGAGACCGAAGACTG 60
DB 1 MethisProThrLeuAsnArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTCGCTCCATCGCGTGGTTCCTTACTCTGCTCAATAACCTCTCAG 120
DB 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGGTATCGCATAGGACAGCTGCACTCCCATAAACCTTATCTCTCACCTGG 180
DB 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGGCACAGGTATTAATATCAACACACTCAAGGGAGGCTCCTTTA 240
DB 61 LeuileThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyAlaProLeu 80
QY 241 GGAACTCGTGGCTGATCTACGTTTCCTTCAGATCAGTTCATTCCTAGTCTGACCTCA 300
DB 81 GlyThrTrpProAspLeuTrpValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCGAGATATCCTCCATGCTCAGGATTTTATGTTTCCCGAGGACCCCAATAATGGA 360
DB 101 ProProAspIleLeuHisAlaHisGlyPheTyValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGGAATCCAGAGATTTCTTTGTAAACATGGAAGTGTAACTCTTAAT 420
DB 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATGGAATGCGCAACCTCTCAGCAGGATAGGTAAGTTTCTTATCTCAAC 480
DB 141 AspGlyTyTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyValAsn 160
QY 481 ACCTATACAGCTCTGGCAATTTTAACTTACCTGACCTCGATTTAGAACGCCCAAG 540
DB 161 ThrTyThrSerSerGlyGlnPheAsnTyLeuThrTrpIleArgThrGlySerProLys 180

QY 541 TGCTCTCTTCAGACCTAGATTACTTAAAAATAAGTTTCTACTGAGAAAGGAAAAACAAGAA 600
DB CysSerProSerAspLeuAspTyLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTCTAAAAATGGGTAATGGTATCTCTTGGGGAATGGTATATTATATGAGGCTCGGT 660
DB AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyTrpTyGlyGlySerGly 220
QY 661 AAACACACAGCTCCATCTTAATTCGGCTCAAAATAAACACAGCTGGAGCCTCCCAATG 720
DB LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGACCAATAATACGGTCTTCACCGGTCAAAAGACCCCAACCCCAAGACACAGGACCA 780
DB 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTTCAACATACTCTGGATCAGACCCCACTGAGTCTAACACAGACACTAAATGGGG 840
DB SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
QY 841 GCAAACTTTTACCTCATCCAGGAGCTTTCAAGCTCTTAAGTCTTCAAGTCTCAGACTCCAGAG 900
DB 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGln 300
QY 901 GCTACCTCTCTCTGCTGCTATGCTTAGCTTCGGGGCCACCTTACTATGAAGGAATGGCT 960
DB 301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyTrpTyGlyMetAla 320
QY 961 AGAAGAGGGAATTCATATGTGACAAAAACATAGACCAATGCATGATGGGGATCCCAA 1020
DB 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAGCTTACCTTACTGAGTTCCTGGAAAGGCGACCTGCATAGGAAAGGTTCCCCCA 1080
DB 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACCAACCTTTGTAAACCACTCAAGACCTTTTAATCAAACTCTGAGAGTCAATAT 1140
DB 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTy 380
QY 1141 CTGCTACTCTGTTATGACAGGTGTGGGCATGTATTAATCTGGAATTAACCCCTTGTGTTCC 1200
DB 381 LeuValProGlyTyTrpAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTAAACCAACTAAAGATTTTGGCAATTTGGTCCAAATGTTTCCCGAGTG 1260
DB 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAAGCAATCTTGATGAATATGACTACAGAAATCATCGACAAAG 1320
DB 421 TyTrpTyTrpProGluLysAlaIleLeuAspGluTyTrpAspTyTrpArgAsnHisArgGlnLys 440
QY 1321 AGAGAACCCATATCTCTGACACTTGTGTGATCTGGACTTGGAGTGGCAGCGGTGTA 1380
DB 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGAACAGCTGCCCTGTCACGGGACCAAGACAGCTAGAAACAGGCTTAGTAAC 1440
DB 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTTGTAAACAGAAAGATCTCCAGCCCTAGAAAAATCTGTCTAGTAACCTGGAG 1500
DB 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACCTCTTATCTGAAGTAGTCTCTACAGAAATAGAGAGGTTAGATTATTA 1560
DB 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGATTTATGTAGCTTGTAGGAGGAGTAATCTGTTTATGTGTGAT 1620
DB 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyTrpValAsp 540
QY 1621 CATTGAGGGGCCATCAGAGACTCCATGAAACAAGCTTAGAGAAAGGTTGGAGAGCGTGA 1680

Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
Qy 1691 AGGGAAGAAAGAACTACTCAAGGGTGGTTTGAGGGATGGTTCAACAGAGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
Qy 1741 GCTACCTACTTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTCTTACTCACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
Qy 1801 GGGCATGCTATTATTAACAAGTTAATTCCTTCAATAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgGileSerAlaValGln 620
Qy 1861 ATCATGTTACTTAGACACAGTACCAGAGCCCGTCTAGCAGGGAAGCTGGCGGC 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 2

AAB70633
ID AAB70633 standard; protein; 638 AA.
AC AAB70633;
XX
DT 15-MAY-2001 (first entry)
DE Porcine endogenous retrovirus envelope protein POBWSL.

KW Porcine endogenous retrovirus; PoERV; gag; env; detection; antigenic;
KW immunogenic; antiviral; vaccine; antiserum; viral infection.
XX
OS Porcine endogenous retrovirus.
XX
PN WO200112816-A2.
XX
PD 22-FEB-2001.

XX 16-AUG-2000; 2000WO-GB003159.
XX
XX 18-AUG-1999; 99GB-00019604.
XX
PA (QUIP-) QUIP TECHNOLOGY LTD.
XX
PI Galbraith D, Kelly H, Smith K;
XX
XX WPI; 2001-211223/21.

XX New porcine endogenous retrovirus polypeptide fragment with retroviral
PT specific antigenic or immunogenic activity, for detection of retroviral
PT antibodies in a sample and in therapy or diagnosis.
XX
PS Claim 7; Fig 2; 43pp; English.

XX The present invention describes a porcine endogenous retrovirus (PoERV)
CC polypeptide fragment (I) which has PoERV specific antigenic or
CC immunogenic activity. Also described are: (1) an antiserum (II) specific
CC to (I); (2) a PoERV specific antibody (III) or its fragment raised
CC against (I); (3) use of (I) in detecting (III) in a sample; (4) use of
CC (III) in detecting PoERV in a sample; (5) use of (I) or (III) in therapy
CC or diagnosis; and (6) an assay kit (IV) comprising (I) or (III) for use
CC in detection of PoERV in a sample. (I) has antiviral activity, and can be
CC used in vaccine production. (I) is useful in the detection of PoERV
CC antibodies (III) in a sample and (III) is useful in the detection of
CC PoERV in a sample. (I) and (III) are useful in therapy or diagnosis.
CC Antiserum (II) to (I) and kits (IV) comprising (I) or (III) are useful in
CC detection of specific PoERV types. Viral infection can be monitored and
CC by using (I) or (III). The present sequence represents a specifically
CC claimed PoERV envelope (env) protein sequence from the present invention
XX
SQ Sequence 638 AA;

Alignment Scores:

Pred. No.:		2.8e-291		Length:		638	
Score:		3425.00		Matches:		638	
Percent Similarity:		100.0%		Conservative:		0	
Best Local Similarity:		100.0%		Mismatches:		0	
Query Match:		98.8%		Indels:		0	
DB:		4		Gaps:		0	
US-10-723-552-3_COPY_5620_7533 (1-1914) x AAB70633 (1-638)							
QY	1	ATGCATCCCAACGTTAAACCGCGCCACCTCCCGATTCGGGTGGAAAGCCGAAAGACTG	60				
DB	1	MethHisProThrLeuAsnArgHisLeuProIleArgGlyGlyLysProLysArgLeu	20				
QY	61	AAAAATCCCTTAAAGCTTCGCCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG	120				
DB	21	LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln	40				
QY	121	ACTAATGGTATGCGCATAGGAGACAGCTGAACTCCCATAAACCCCTTATCTCTCACCTGG	180				
DB	41	ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp	60				
QY	181	TTAATTACTGACTCCGGCACAGGTATTAAATATCAACAACTCAAGGGAGGCTCTTTA	240				
DB	61	LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu	80				
QY	241	GGAACTGGTGGCTGATCTATACGTTTTCCTCAGATCAGTTATTTCTAGTCTGACCTCA	300				
DB	81	GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer	100				
QY	301	CCCCAGATATCTCCATGCTCAGGATTTTATGTTTCCCGAGACCAACCAATAATGGA	360				
DB	101	ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly	120				
QY	361	AAACATTCGCGAAATCCCAAGAGATTTCTTTTGTAAACAATGGAACCTGTAAACCTCTAAT	420				
DB	121	LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn	140				
QY	421	GATGATATTGGAAATGGCCAACTCTCAGCAGATAGGGTAACTTTTCTTATGTCAAC	480				
DB	141	AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn	160				
QY	481	ACCTATACAGCTCTGGACAAATTAATCTGACCTGGATAGAACTGGAAGCCCCAAG	540				
DB	161	ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys	180				
QY	541	TGCTCTCCTTCAGACCTAGATTACCTAAAAATAAGTTTCACTGAGAAGGAAACAGAA	600				
DB	181	CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu	200				
QY	601	AATATCCTAAATAGGGTAAATGGTATGCTCTTGGGAATGGTATATTATGAGGCTCGGGT	660				
DB	201	AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGly	220				
QY	661	AAACAAACAGGCTCCATTTCACTATTTCGCTCAAAATAAACACAGCTGGAGCTCCAATG	720				
DB	221	LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet	240				
QY	721	GCTATAGGACCAATATACGCTCTTGACGGGTCAAGACCCCAACCCCAAGGACGAGGACCA	780				
DB	241	AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro	260				
QY	781	TCCTCTAAATAACTTCTGGATCAGACCCCACTGAGTCTAACACAGCAGCTAAATATGGG	840				
DB	261	SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly	280				
QY	841	GCAAACTTTTATAGCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAG	900				
DB	281	AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu	300				
QY	901	GCTACTCTTCTTGTGGCTATGCTTAGCTTCGGGCCCACTTACTATGAGGAAATGGCT	960				
DB	301	AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMetAla	320				

QY 361 AAACATTGCGGAATCCAGAGATTCTTTTGTAAACAAATGGAACCTGTGTAACTCTAAT 420
Db |||||
121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCyseValThrSerAsn 140
QY 421 GATGGATATTGGAATGCCACCTCTCAGCAGGATAGGTAAGTTTCTTATGTCAAC 480
Db |||||
141 AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACAGCTCTGGACAATTTAAATTACCTCGACTGATTTAGAACTGGAGGCCCAAG 540
Db |||||
161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCCTTCAGACCTAGATTACCTAAATAAAGTTTCACTGAGAAAGGAAACAAGAA 600
Db |||||
181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTCTAAATGGTAAATGATGCTCTGGGGAATGGTATATTATGAGAGGCTCGGT 660
Db |||||
201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyLysGly 220
QY 661 AAACAACAGAGCTCCATTCTAACTATTCCGCTCAAAATAAACACAGCTGGAGCCTCCAATG 720
Db |||||
221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProPheMet 240
QY 721 GCTATAGGACCAATACGGTCTTGACGGGTCAAGACCCCAAGGACCCAGGACCA 780
Db |||||
241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTCTCGATCAGACCCCACTGAGTCTAACAGACGACTAAATGGGG 840
Db |||||
261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
QY 841 GCAAACTTTTAGCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCACGACTCCAGAG 900
Db |||||
281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTCTCTGTGGCTAGCTTAGCTTCGGGCCCCACCTTACTATGAGGAATGGCT 960
Db |||||
301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMetAla 320
QY 961 AGAAGAGGAAATTCATGTGACAAAAGAACATGAGACCAATGCACATGGGGATCCAA 1020
Db |||||
321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGTCTTCTGGAAAGCGACCTGCATAGAAAGGTTCCCCCA 1080
Db |||||
341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCACCAACACCTTTGTAACCACTGAAGCCTTTAATCAAACTCTGAGAGTCAATAT 1140
Db |||||
361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGTACTCTGTTATGACAGTGTGGCATGTAACTAGTAAACCCCTTGTGTTTCC 1200
Db |||||
381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGTGTTTTAACCAACTAAAGATTTTGTCAATTATGTCCTCAAAATGTTCCCGAGTG 1260
Db |||||
401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAAGCAATCCTTGTATGAATATGACTACAGAAATCATCCGACAAAAG 1320
Db |||||
421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAAGACCCATATCTTGACACTTGTGATGTCGGACTTGGAGTGGCAGCAGGTGTA 1380
Db |||||
441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGACAGCTGCCCTGCTCAGCGGACCCAGCAGCTAGAAACAGGACTTAGTAAC 1440
Db |||||
461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480

QY 1441 CTACATCGAATTGTAAACAGAGATCTCCAAGCCCTAGAAAAATCTGTCAAGTAACCTGGAG 1500
Db |||||
481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACCTCTTATCTGAAGTAGTCTCAGAGATAGAGAGGGTTAGATTATTA 1560
Db |||||
501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGATTATGTAGCCCTTGAAGGAGGAATGCTGTTTTTATCTGGAT 1620
Db |||||
521 PheLeuLysGluGlyGlyLeuCyseValAlaLeuLysGluGluCysCysePheTyrValAsp 540
QY 1621 CATTCAAGGGCCCATCAGAGACTCCATGAACAAGCTTAGAGAAAAGTTTGGAGAAGCGTCCA 1680
Db |||||
541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGGAAGAGAACTACTCAAGGCTGTTTCAGGGATGGTTCAACAGGCTCTTTGGTTG 1740
Db |||||
561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
QY 1741 GCTACCTTACTTCTGCTTTTAAACAGGACCCCTTAATAGTCTCTCTCTTACTCACAGTT 1800
Db |||||
581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATGTATTATTAAACAAGTTAATTCCTTTCATTAGAGAACGAATAAGTCAGTCCAG 1860
Db |||||
601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGTACTTAGACACACAGTACCAAGCCCTCTAGCAGGGAAGCTGGCCGC 1914
Db |||||
621 IleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 638

RESULT 4
ADS73448
ID ADS73448 standard; protein; 638 AA.
XX ADS73448;
XX AC
XX DT 16-DEC-2004 (first entry)
XX DE Swine retroviral env protein.
XX KW Swine retroviral protein; immunosuppressive; gene therapy; env protein.
XX OS Porcine endogenous retrovirus.
XX PN US2004185435-A1.
XX PD 23-SEP-2004.
XX PF 26-NOV-2003; 2003US-00723552.
XX PR 14-DEC-1995; 95US-00572645.
XX PR 13-DEC-1996; 96US-00766528.
XX PR 14-SEP-2000; 2000US-00661858.
XX PA (GENO) GEN HOSPITAL CORP.
XX PI Fishman JA;
XX DR WPI; 2004-689179/67.
XX DR N-PSDB; ADS73369.
XX PT New porcine retroviral polypeptide encoded by a nucleic acid, useful in
XX PT evaluating an immunosuppressive treatment for the ability to activate a
XX PS retrovirus, such as an endogenous porcine retrovirus.
XX PS Claim 8; Fig 3; 83pp; English.
XX CC The present invention relates to the swine retroviral polypeptides and
CC their encoding nucleic acids. The methods and compositions of the present
CC invention are useful for screening a cell or tissue, e.g. a heart, lung,
CC liver, bone marrow, kidney, brain cells, neural tissue, pancreas and

CC intestinal tissue xenograft, for the presence or expression of a swine or
CC miniature swine retrovirus or retroviral sequence. The invention is also
CC useful in evaluating an immunosuppressive treatment for the ability to
CC activate a retrovirus such as an endogenous porcine retrovirus. The
CC invention is also useful in gene therapy. The present sequence is the
CC swine retroviral env protein.
XX
SQ Sequence 638 AA;

Alignment Scores:

Pred. No.: 2,8e-291 Length: 638
Score: 3425.00 Matches: 638
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 98.8% Indels: 0
DB: 8 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x ADS73448 (1-638)

```
QY 1 ATGCATCCACGCTTAACCGCGGCACCTCCCGATTCCGGGTGGAAAGCCGAAAGACTG 60
DB 1 MethisProThrLeuAsnArgArgHisLeuProIleArgGlyGlyLeuProLysArgLeu 20
QY 61 AAAATCCCTTAAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG 120
DB 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGGTATGCGCATAGGAGACGCTGAACCTCCCATAAACCTTATCTCTCACCTGG 180
DB 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCGGCACAGGTATTAAATACACACACTCAAGGGAGGCTCCCTTA 240
DB 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTCGTGGCTGATCTATACGTTTGCCTCAGATCAGTATTATCTCTAGTCTGACCTCA 300
DB 81 GlyThrTrpTrpProAspLeuTyValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCTCCATGCTCAGGATTTATGTTTCCAGGACCCAGCACCAAAATAAGGA 360
DB 101 ProProAspIleLeuHisAlaHisGlyPheTyValCysProGlyProProAsnAsnGly 120
QY 361 AAACNTTCGGAAATCCAGAGATTCTTTTGTAAACATGGAATGTAACCTCTAAT 420
DB 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTGGAATGGCCAACTCTCAGCAGGATAGGTAAGTTTCTTATGTCAAC 480
DB 141 AspGlyTyTrpLysTrpProThrSerGlnAspArgValSerPheSerTyValAsn 160
QY 481 ACCTATACCAGCTCTGGCAATTTAATTTACCTGACCTCGATAGAACTGGAGCCCCAAG 540
DB 161 ThrTyThrSerSerGlyGlnPheAsnTyLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCCTTCAGACCTAGATTACCTAAATAAGTTTCACCTGAGAAAGGAAACAGAA 600
DB 181 CysSerProSerAspLeuAspTyLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTTAAATGGTAAATGGTATGTTCTTGGGAATGGTATATTATGGAGGCTCGGGT 660
DB 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyTrpGlyLysSerGly 220
QY 661 AAACACAGGCTCCATTTCTAATTCGCTCAAAATAAACCAGCTCGAGCCTCCAATG 720
DB 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGACCAATACCGTCTTGACGGTCAAGACCCCAACCCAGGAGGACGACCA 780
DB 241 AlaIleGlyProAsnThrValIleuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTTCTGGATTCAGACCCCACTGAGTCTTAACAGCACGACTAAATATGGG 840
```

```
DB 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
QY 841 GCRAAATCTTTTACCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCACAGCTCCAGAG 900
DB 281 AlalysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAenSerThrThrProGlu 300
QY 901 GCTACCTCTTCTTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGGAATGGCT 960
DB 301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyTrpTrpGluGlyMetAla 320
QY 961 AGAAGAGGAAATCAATGTGACAAAAACATAGACCAATAGACCAATGATGGGATCCCAA 1020
DB 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTGGRAAAGGACCTGCATAGGAAGGTTCCCCCA 1080
DB 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCACCAACACCTTTGTAAACCACTGAAGCCCTTTAATCAAACTCTGAGAGTCAATAT 1140
DB 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTy 380
QY 1141 CTGTACTCTGTTATGACAGTGTGGGCATGTAATACTGGATTAAACCCCTGTGTTC 1200
DB 381 LeuValProGlyTyAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTAAACCAACTAAAGATTTTGTGATTATGTCCTCAAAATGTTCCCGAGTG 1260
DB 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAAGCAATCCTTGATGAATATGACTACAGAAATATCATCGCAAAAG 1320
DB 421 TyTrpTyTrpProGluLysAlaIleLeuAspGluTyAspTyArgAsnHisArgGlnLys 440
QY 1321 AGAAGAACCATATCTCTGACACTGCTGTGATGCTCGGACTTGGAGTGGCAGAGGTGTA 1380
DB 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGACAGCTGCCCTGTCACGGGACCCAGCAGCTAGAACAGGACTTAGTAAC 1440
DB 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTTAAACAGAGATCTCAAGCCCTAGAAAAATCTGTCTAGTAACCTCGAG 1500
DB 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACCTCTTATCTGAAGTAGTCTCTACAGATAGAGAGGTTAGATTTATA 1560
DB 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGATTTATGTAGCCCTTGAAGGAGGAATGCTGTTTTTATGTGGAT 1620
DB 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyValAsp 540
QY 1621 CATTCAGGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAAGGTTTCGAGAAAGCGTGA 1680
DB 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGGAAGAGGAACTACTCAAGGGTGGTTTTCAGGGATGGTTCAACAGGCTCTTTGGTTG 1740
DB 561 ArgGluLysGluThrThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
QY 1741 GCTACCTCTCTTCTGCTTTTAAACAGGACCTTAACTAGTCTCTCCCTCTGTTACTCACAT 1800
DB 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCATGTATTATTAAACAGTTAATTCCTCTTCAATAGAGAACCAATTAAGTCAGTCCAG 1860
DB 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgGlyIleSerAlaValGln 620
QY 1861 ATCATGTGCTTACAGAACAGTACCAAGCCGCTCTAGCAGGAGAGCTGGCCGC 1914
DB 621 IleMetValLeuArgGlnGlnTyArgSerProSerArgGluAlaGlyArg 638
```



```
RESULT 5
AAB35114
ID AAB35114 standard; protein; 638 AA.
XX
AC AAB35114;
XX
DT 27-MAR-2001 (first entry)
XX
DE PERV-C env protein SEQ ID NO: 5.
XX
KW Xenotransplantation; infectious agent; vaccine.
XX
OS Porcine endogenous retrovirus.
XX
PN WO200071726-A1.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014296.
XX
PR 24-MAY-1999; 99US-0135631P.
XX
PA (MAYO-) MAYO MEDICAL VENTURES.
XX
PI Federspiel MJ;
XX
DR WPI; 2001-032041/04.
XX
PT Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells.
XX
PS Disclosure; Page 95-97; 144pp; English.
XX
CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient
XX
SQ Sequence 638 AA;
Alignment Scores:
Pred. No.: 7,69e-291 Length: 638
Score: 3420.00 Matches: 637
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 98.6% Indels: 0
DB: 4 Gaps: 0
US-10-723-552-3_COPY_5620_7533 (1-1914) x AAB35114 (1-638)
QY 1 ATGCATCCACGTTAAACCGCGCCACCTCCCGATTCGGGTGGAAAGCGGAAAGACTG 60
DB 1 MethisProThrLeuAsnArgAgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTTCCTTACTCTGTCAATAACCTCTCAG 120
DB 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATCGGTATGCGCATAGGAGACAGCTCGAATCCCAATACCTTATCTCTCACTGG 180
DB 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAAGTACCTCCGACAGGATTAATATCAACAACACTCAAGGGGAGGCTCCCTTTA 240
DB 61 LeuileThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATAGCTTGGCTCAGATCAGTTATTCCTAGCTCAGCTCA 300
DB 81 GlyThrTrpTrpProAspLeuThrValCysLeuArgSerValIleProSerLeuThrSer 100
```

```
QY 301 CCCCAGATATCTCTCCATGCTCACGGATTTTATGTTTCCAGGACCAACCAATAATGGA 360
DB 101 ProProAspIleLeuHisAlaHisGlyPheTyValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTTGGGAAATCCAGAGATTTCTTTGTTAAACAATGAACTGTGTAACTCTTAAT 420
DB 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGGATATTGGAAATGCCAACCTCTCAGCAGATAGGGTAACTTTTCTTATGTCAC 480
DB 141 AspGlyTyTrpLysTrpProThrSerGlnAspArgValSerPheSerTyValAsn 160
QY 481 ACCTATACAGCTCTGCACAAATTTAAATTACTGACCTGGATTAGAACTGGAGGCCCAAG 540
DB 161 ThrTyThrSerSerGlyGlnPheAsnTyLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCTTCAGACCTAGATTACCTAAATAAGTTTCACTGAGAAGGAAAAACAAGAA 600
DB 181 CysSerProSerAspLeuAspTyLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCCTAAATGGGTAAATGGTATGCTCTGGGAATGGTATATTTATGAGGCTCGGT 660
DB 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyTrpGlyGlySerGly 220
QY 661 AAACAACACGAGCTCCATTCTAACTATTTCGCTCAAAATAAACACAGCTGGAGCCTCCAATG 720
DB 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProPmet 240
QY 721 GCTATAGGACCAATACGGTCTTGACGGGTCAAGACCCCAACCCCAAGGACCAAGACCA 780
DB 241 AlaileGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTCTCGATCAGACCCCACTGAGCTTAACACAGCAGACTAAATGGG 840
DB 261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerSerThrThrLysMetGly 280
QY 841 GCAAAACCTTTTACGCTCATCCAGGGAGCTTTTCAAGCTCTTAAGTCTCAAGCTCCAGAG 900
DB 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTTTCTTGGCTATGCTTAGCTTCGGGCCCACTTACTATGAAGGAATCGCT 960
DB 301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyTrpGlyMetAla 320
QY 961 AGAAGAGGGAATTCATGTGACAAAAGACATAGAGACCAATGCACATGGGATCCCAA 1020
DB 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGGTTCCTGGAAAAAGGACCTGCATAGGAAAGTTCCCCCA 1080
DB 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACCAACACCTTTGTAAACACACTGAAGCCCTTTAATCAAACTCTGAGAGTCAATAT 1140
DB 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTy 380
QY 1141 CTGGTACCTGGTTATCAGAGGTGGGTGGATTAATCTGGAATTAACCCCTTGTGTTTCC 1200
DB 381 LeuValProGlyTyTrpAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTCTGTTTAAACCAAACTAAAGATTTTTCATTTATGTCCTCAAAATGTTTCCCGAGTG 1260
DB 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCCGAAAAAGCAATCTCTTGATGATATGACTACAGAAATCAACGAAAAAG 1320
DB 421 TyTrpTyTrpProGlyLysAlaIleLeuAspGluTyTrpAspTyTrpArgAsnHisArgGlnLys 440
QY 1321 AGAAGAACCATATCTCTGACACTTGTGTGATGCTCGGACTCGGAGTGGAGTGGCAGGTTGA 1380
DB 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
```

QY 1381 GGAACAGGACAGCTGCTGCTACGGGACACACAGCAGCTAGAACAGGACTTAGTAAC 1440
Db |||||||
QY 461 GlyThrGlyThrAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
Db |||||||
QY 1441 CTACATCGAATTGTAACGAAGACTCTCCAAGCCCTAGAAAAATCTGTCTAGTAACCTGGAG 1500
Db |||||||
QY 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTAACCTCTTATCTGAAGTAGTCTCTACAGATAGAGAAGGGTTAGATTATTA 1560
Db |||||||
QY 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGAGGATTATGTGTAGCTTTGAAGGAGGAATGCTGTTTTATCTGCAT 1620
Db |||||||
QY 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysPheTyrValAsp 540
QY 1621 CATTGAGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAGGTTGGAGAAGCGTCCA 1680
Db |||||||
QY 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGNAAGAAAGAACTACTCAAGGTGGTGTGAGGATGGTTCACAGCTCTCTTGGTTG 1740
Db |||||||
QY 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
QY 1741 GCTACCTACTTCTGCTTTAACAGGACCCCTAATAGTCTCTCTCTCTCTCTCTCTCT 1800
Db |||||||
QY 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCATGTATTATTAAACAAGTTAATTCCTTCATTAGAGAACGAATTAAGTCAGTCCAG 1860
Db |||||||
QY 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGGTACTTAGACACAGTACCAAGCCGCTCTAGCAGGAAGCTGGCGCGC 1914
Db |||||||
QY 621 IleMetValLeuArgGlnGlnTrpGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 6
ADY28028

ID ADY28028 standard; protein; 638 AA.
XX
AC ADY28028;
XX
DT 19-MAY-2005 (first entry)
XX
DE Porcine endogenous retrovirus subtype C ancestral env protein, SEQ ID 32.
XX

KW Xenotransplantation; vaccine; acquired immune deficiency syndrome;
KW anti-hiv; immune disorder; HIV-1 infection; infection; PERV infection;
KW diagnosis; graft versus host disease; immunosuppressive;
KW envelope protein.

XX Porcine endogenous retrovirus; subtype C.
OS Synthetic.

XX WO2005019411-A2.

XX 03-MAR-2005.

XX 19-MAY-2004; 2004WO-US015709.

XX 19-MAY-2003; 2003US-00441949.

XX (AUCK-) AUCKLAND UNISERVICES LTD.
PA (UNITW) UNIV WASHINGTON.

XX Mullins JI, Rodrigo A, Ross HA;

XX WPI; 2005-202625/21.

DR N-PSDB; ADY28027.

PT New isolated ancestral viral nucleic acid sequence that is a determined
PT founder sequence of a highly diverse viral strain, subtype or group of an
PT endogenous retrovirus, useful in the prophylaxis of viral infection.

XX Claim 10; SEQ ID NO 32; 278pp; English.
PS
XX The invention relates to an isolated ancestral viral nucleic acid
CC sequence and its fragment, where the sequence is a determined founder
CC sequence of a highly diverse viral strain, subtype or group of an
CC endogenous retrovirus. Also included are an isolated ancestor protein (or
CC its fragment) from an endogenous retrovirus, an isolated expression
CC construct (comprising the following operably linked elements:- a
CC transcriptional promoter; a nucleic acid encoding an endogenous
CC retrovirus ancestor protein and a transcriptional terminator), a cultured
CC prokaryotic (or eukaryotic cell) transformed or transfected with the
CC expression construct, an isolated host cell comprising the expression
CC construct, a composition for inducing an immune response in a recipient
CC mammal (comprising a viral ancestor protein or its antigenic fragment,
CC where the protein is from a virus of a donor species), an isolated
CC antibody that binds specifically to an endogenous retrovirus ancestor
CC protein (and that binds specifically to a plurality of circulating
CC descendant endogenous retrovirus ancestor proteins), a method of
CC preparing an ancestral endogenous retroviral amino acid sequence, a
CC method for inducing an immune response to a donor virus in a transplant
CC recipient or a potential transplant recipient, a method of making a
CC vaccine, a method for detecting infection with an endogenous retrovirus
CC and a method for performing xenotransplantation in a subject. The
CC ancestral viral nucleic acid sequence is of Porcine Endogenous Retrovirus
CC (PERV) subtype A, B or C and has at least 70% identity with sequence of
CC any of the 18 nucleotide sequences fully defined in the specification.
CC The sequence may also be optimized for expression in a human host. The
CC nucleic acid sequence or its fragment is useful in the prophylaxis of
CC viral infection in transplantation that is heightened by the presence of
CC factors commonly associated with viral activation, e.g. immune
CC suppression, graft versus host disease, graft rejection, viral co-
CC infection, and cytotoxic therapies. The present sequence is an ancestral
CC PERV env protein.
XX
SQ Sequence 638 AA;

Alignment Scores:
Pred. No.: 1,08e-289 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x ADY28028 (1-638)

QY 1 ATGCATCCCGTAAACCCGCGCCACCTCCCGATTCCGGGTGGAAGCCGAAAGACTG 60
Db 1 Methi.sProThrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAAGCTTCGCCTCCATCGCGTGGTTCCTTACTCTGTCAATACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGCGTAGCCGATAGGACAGCAGCTGAACCTCCATAAACCTTATCTCTACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGGCAGCAGGTATTAAATATCAACACACTCAAGGGAGGCTCTTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyLeuAsnIleAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCCTGATCTATAGCTTTGCCCTCAGATCAGTTATCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCGAGGACCAACCAATAATCGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGCGAAATCCGAGAGATTCTTTTGTAAACAAATGGAACGTGTGAACCTCTAAT 420
Db |||||||

Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTGGAATGCCCAACCTCTCAGCAGGATAGGTAAGTTTTCTTATGTCAAC 480
Db 141 AspGlyTyrTrpLysTrpThrSerGlnGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACAGCTCTGGCAATTTAATTACCTGACCTGGATTAGAACTGGAGCCCAAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCTCAGACCTAGATTACCTAAATAAAGTTTCACTGAGAAAGGAAACAGAA 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGln 200
QY 601 AATATCCATAAAGTGGTAAATGGTATGTCTTGGGAATGGTATATTATGAGGCTCGGGT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGly 220
QY 661 AAACACACAGGCTCAATCTTAACCTATTCGCCTCAAAATAAACACAGCTGGAGCCCTCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProMet 240
QY 721 GCTATAGGACCAATACCGCTCTCAGCGGTCAAGACCCCAACCCAGGACGAGGACCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAAACATAACTCTCGGATCAGACCCCACTGAGTCTAAGCTTAACACGACGACTAAATGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
QY 841 GCAAACTTTTACGCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCACGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTTCTTGTGGTATGCTTACCTTGTGGGCCACCTTACTATGAGGAATGGCT 960
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGlyMetAla 320
QY 961 AGAAGAGGGAATCAATGTGACAAAGAACATAGAGACCAATGCATGGGGATCCAA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACGAGTTTCTGAAAGGACCTGCATAGAAAGGTTTCCCCCA 1080
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCACCAACACCTTGTACCACTGAAGCTTTAATCAACCTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGGTACCTGGTTATGACAGGTGGGCGCATGTAACTGATTAACCCCTTGTGTTTCC 1200
Db 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTAAACCAAACTAAAGATTTTTCATATTGTCCTCAAAATGTTTCCCCGAGTG 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGCAAAAG 1320
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAAGACCCATATCTCTGACACTTGTGTGTATGCTCGGACTTGGAGTGGCAGGAGGTA 1380
Db 441 ArgGluProLysSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGGAACAGCTGCCCTGGTCACGGGACCAAGCAGCTAGAAAACAGGACTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTTGTAACAGAGATCTCCAGGCCCTAGAAAATCTGTCAGTAACCTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500

QY 1501 GAATCCCTAACCTCTTATCTGAAGTAGTCTCCTACAGAAATAGAGGGTTAGATTATTA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAAGAGGAGGATTATGTAGCTTACCTTGAAGGAGGAATGCTGTTTTATGTGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGlyCysPheTyrValAsp 540
QY 1621 CATTCAAGGGGCATCAGAGACTCCATCAAGCAAGCTTAGAGAAAGGTTGGAGAAGCTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluLeuLysArgArg 560
QY 1681 AGGAAAAGGAAACTACTCAAGGCTGTTTGAAGGATGGTTCAACAGGCTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
QY 1741 GCTACCTACTTCTTCTTAAACAGGACCTTAATAGTCTCTCTCTCTCTCTCTCTCT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATGTATTATTAACAAGTTAATTCCTTCAATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGCTACTTAGACAACTAGTACCAAGCCCGTCTAGCAGGGAAGCTGGCCGC 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 7

ADY28030

ID ADY28030 standard; protein; 638 AA.

AC ADY28030;

DT 19-MAY-2005 (first entry)

XX Porcine endogenous retrovirus subtype C ancestral env protein, SEQ ID 34.
XX Xenotransplantation; vaccine; acquired immune deficiency syndrome;
XX anti-hiv; immune disorder; HIV-1 infection; infection; PERV infection;
XX diagnosis; graft versus host disease; immunosuppressive;
XX envelope protein.

OS Porcine endogenous retrovirus; subtype C.
OS Synthetic.

PN WO2005019411-A2.

PD 03-MAR-2005.

XX 19-MAY-2004; 2004WO-US015709.

XX 19-MAY-2003; 2003US-00441949.

XX (AUCC-) AUCKLAND UNISERVICES LTD.
XX (UNIW) UNIV WASHINGTON.

PI Mullins JI, Rodrigo A, Ross HA;

XX WPI; 2005-202625/21.

DR N-PSDB; ADY28029.

XX New isolated ancestral viral nucleic acid sequence that is a determined
XX founder sequence of a highly diverse viral strain, subtype or group of an
XX endogenous retrovirus, useful in the prophylaxis of viral infection.

PS Claim 10; SEQ ID NO 34; 278pp; English.

XX The invention relates to an isolated ancestral viral nucleic acid
XX sequence and its fragment, where the sequence is a determined founder
XX sequence of a highly diverse viral strain, subtype or group of an
XX endogenous retrovirus. Also included are an isolated ancestor protein (or

CC its fragment) from an endogenous retrovirus, an isolated expression
CC construct (comprising the following operably linked elements:- a
CC transcriptional promoter; a nucleic acid encoding an endogenous
CC retrovirus ancestor protein and a transcriptional terminator), a cultured
CC prokaryotic (or eukaryotic cell) transformed or transfectected with the
CC expression construct, an isolated host cell comprising the expression
CC construct, a composition for inducing an immune response in a recipient
CC mammal (comprising a viral ancestor protein or its antigenic fragment,
CC where the protein is from a virus of a donor species), an isolated
CC antibody that binds specifically to an endogenous retrovirus ancestor
CC protein (and that binds specifically to a plurality of circulating
CC descendant endogenous retrovirus ancestor proteins), a method of
CC preparing an ancestral endogenous retroviral amino acid sequence, a
CC method for inducing an immune response to a donor virus in a transplant
CC recipient or a potential transplant recipient, a method of making a
CC vaccine, a method for detecting infection with an endogenous retrovirus
CC and a method for performing xenotransplantation in a subject. The
CC ancestral viral nucleic acid sequence is of Porcine Endogenous Retrovirus
CC (PERV) subtype A, B or C and has at least 70% identity with sequence of
CC any of the 18 nucleotide sequences fully defined in the specification.
CC The sequence may also be optimized for expression in a human host. The
CC nucleic acid sequence or its fragment is useful in the prophylaxis of
CC viral infection in transplantation that is heightened by the presence of
CC factors commonly associated with viral activation, e.g. immune
CC suppression, graft versus host disease, graft rejection, viral co-
CC infection, and cytotoxic therapies. The present sequence is an ancestral
CC PERV env protein.
XX

SQ Sequence 638 AA;

Alignment Scores:
Pred. No.: 1,06e-289 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: 9 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x ADV28030 (1-638)

QY 1 ATGCATCCAGCTTAAACCGCGCCACCTCCGATTCGGGTGGGAAAGCCGAAAGACTG 60
DB 1 MethHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAAGCTTCGCTCCATCGCGGTTCCTTACTCTGTCAATTAACCTCTCAG 120
DB 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGCTATGCGCATAGGACAGCCTGAACTCCCATAAACCTTATCTCTCACCTGG 180
DB 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGGCACAGGTATTAAATATCAACACACTCAAGGGGAGCTCTTTA 240
DB 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyAlaProLeu 80
QY 241 GGAACCTGGTGGCTGTATACGTTTGCCTCAGATCAGTATTCTCTAGTCTGACCTCA 300
DB 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCCTCATGCTCACGGATTTATGTTTCCCGAGGACCCACCAATATGGA 360
DB 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGAAATCCAGAGATTCTTTGTAAACAATGGAACTGTGTAACCTCTAAT 420
DB 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTTGGAATGCGCAACCTCTCAGCAGGATAGGGTAAGTTTCTTATGTCAAC 480
DB 141 AspGlyTyrTrpLysTrpProThrSerGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACGAGCTCTGGACAAATTTAATTACCTGACCTGGATTAGAACTGGAAGCCCCAAG 540

DB 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCCTCAGACCTAGATTACCTAAAAATAATAGTTTCTACTGAGAAAGGAAACAAGAA 600
DB 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTTAAATGGGTAAATGGTATGCTCTTGGGGAATGGTATATTATGAGGCTCGGGT 660
DB 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrGlyGlySerGly 220
QY 661 AAACAACAGGCTCCATTTCTAACTATTTCGCTCAAAATAAAACCAAGCTGGAGCTCCCAATG 720
DB 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGACCAATATACGCTTCGACGGGTCAAAAGACCCCAACCCAAAGGACCAAGACCA 780
DB 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATACTTCTGGATCAGCCCACTGAGTCTAACAGCAGCAGCAGCTAAAAATGGG 840
DB 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
QY 841 GCAAACTTTTACCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCACGACTCCAGAG 900
DB 281 AlalysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTTCTTGGCTATGCTTAGCTTCGGGCCACCTTATCTATGAAGGAATGGCT 960
DB 301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAla 320
QY 961 AGAAGAGGGAATTCATGTGCAAAAGAACATAGACCAATCAGATGGGATGCCAA 1020
DB 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTGGAAAGGACCTGTCATAGGAAAGGTTTCCCCCA 1080
DB 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACCAACCTTTGTAAACCACTGAAGCCTTTTAACTCAAACTCTGAGAGTCAATAT 1140
DB 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGCTACTGCTGTTATGACAGTGTGGGCATGTAATACTGGAATTAACCCCTTGTGTTTCC 1200
DB 381 LeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTTTAAACCAAACTTAAAGATTTTTCATTTATGTCCTCAAAATTTTCCCGAGTG 1260
DB 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAG 1320
DB 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGACAACCATATCTCTGACACTTCTGTGATGCTCGGACTTCGGAGTGGAGCGCAGGCTGA 1380
DB 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGGAACAGCTGCCCTCGTCCAGGGACCAACAGCAGCTAGAACAGGACTTAGTAAC 1440
DB 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTTGACAGAAAGATCTCCAAGCCCTAGAAAAATCTGTCTAGTAACCTGAG 1500
DB 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACCTCTTATCTGAAGTAGTCTCAGAGTAGCAGAGTGAAGGGTTAGATTATTA 1560
DB 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAAAGGAGGATTATGTGTAGCCTTTGAAGGAGGAATGCTGTTTTTATGTGAT 1620

Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysPheTyrValAsp 540
Qy 1621 CATTGAGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAAAGTTGGAGAAGCGTGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
Qy 1681 AGGGAAAGGAACACTCAAGGTGGTTTGGAGGATGGTTCAACAGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
Qy 1741 GCTACCCCTACTTCTGCTTTTAACAGGACCCTTAATAGTCCCTCCCTGTTACTCACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
Qy 1801 GGGCCATGATATTAAACAAGTTAATTCCTTCATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
Qy 1861 ATCATGCTACTTAGACACAGTACCAAGCCCGTCTAGCAGGGAAGCTGGCCGC 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 638

RESULT 8

ID ADY28036 standard; protein; 638 AA.
AC ADY28036;
XX
DT 19-MAY-2005 (first entry)
XX
DE Porcine endogenous retrovirus subtype C ancestral env protein, SEQ ID 40.
XX
KW Xenotransplantation; vaccine; acquired immune deficiency syndrome;
KW anti-hiv; immune disorder; HIV-1 infection; infection; PERV infection;
KW diagnosis; graft versus host disease; immunosuppressive;
KW envelope protein.

OS Porcine endogenous retrovirus; subtype C.

OS Synthetic.

PN WO2005019411-A2.

PD 03-MAR-2005.

PF 19-MAY-2004; 2004WO-US015709.

PR 19-MAY-2003; 2003US-00441949.

PA (AUCK-) AUCKLAND UNISERVICES LTD.

PA (UNIW) UNIV WASHINGTON.

PI Mullins JI, Rodrigo A, Ross HA;

DR WPI; 2005-202625/21.

DR N-PSDB; ADY28035.

XX New isolated ancestral viral nucleic acid sequence that is a determined
PT founder sequence of a highly diverse viral strain, subtype or group of an
PT endogenous retrovirus, useful in the prophylaxis of viral infection.

PS Claim 10; SEQ ID NO 40; 278pp; English.

XX The invention relates to an isolated ancestral viral nucleic acid
CC sequence and its fragment, where the sequence is a determined founder
CC sequence of a highly diverse viral strain, subtype or group of an
CC endogenous retrovirus. Also included are an isolated ancestor protein (or
CC its fragment) from an endogenous retrovirus, an isolated expression
CC construct (comprising the following operably linked elements:- a
CC transcriptional promoter; a nucleic acid encoding an endogenous
CC retrovirus ancestor protein and a transcriptional terminator), a cultured
CC prokaryotic (or eukaryotic cell) transformed or transfected with the
CC expression construct, an isolated host cell comprising the expression
CC construct, a composition for inducing an immune response in a recipient

CC mammal (comprising a viral ancestor protein or its antigenic fragment,
CC where the protein is from a virus of a donor species), an isolated
CC antibody that binds specifically to an endogenous retrovirus ancestor
CC protein (and that binds specifically to a plurality of circulating
CC descendant endogenous retrovirus ancestor proteins), a method of
CC preparing an ancestral endogenous retroviral amino acid sequence, a
CC method for inducing an immune response to a donor virus in a transplant
CC recipient or a potential transplant recipient, a method of making a
CC vaccine, a method for detecting infection with an endogenous retrovirus
CC and a method for performing xenotransplantation in a subject. The
CC ancestral viral nucleic acid sequence is of Porcine Endogenous Retrovirus
CC (PERV) subtype A, B or C and has at least 70% identity with sequence of
CC any of the 18 nucleotide sequences fully defined in the specification.
CC The sequence may also be optimized for expression in a human host. The
CC nucleic acid sequence or its fragment is useful in the prophylaxis of
CC viral infection in transplantation that is heightened by the presence of
CC factors commonly associated with viral activation, e.g. immune
CC suppression, graft versus host disease, graft rejection, viral co-
CC infection, and cytotoxic therapies. The present sequence is an ancestral
CC PERV env protein.

XX Sequence 638 AA;

Alignment Scores:

Pred. No.: 1.06e-289 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x ADY28036 (1-638)

Qy 1 ATGCATCCACAGTTAAACCCGGCCACCTCCCGATTCGGGTGAAAGCGAAAGACTG 60
Db 1 MethisProThrLeuSerArgHiseuProIleArgGlyGlyLysProLysArgLeu 20
Qy 61 AAAATCCCTTAAGCTTCGCCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
Qy 121 ACTAATGGTATGGCATAGGAGACAGCTGAACCTCCCATAAACCTTATCTCTCACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
Qy 181 TTAATTACTGACTCCGSCACAGGTATTAAATATCAACAACACTCAAGGGGAGCTCCTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
Qy 241 GGAACTGGTGGCTGATCTATACGTTTGCTCAGATCAGTATTATCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
Qy 301 CCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGACCAACAAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
Qy 361 AAACATTGCGGAATCCACAGATTTCTTTTGTAAACAATGGAACCTGTAACTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
Qy 421 GATGGATATTGGAAATGGCCAACCTCTCAGCAGATAGGGTAAAGTTTCTTATGTCAAC 480
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnAspArgValSerPheSerTyrValAsn 160
Qy 481 ACCTATACAGCTCTGGACAATTTAATTACTGACCTGGATTAGAACTGGAAGCCCCAAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
Qy 541 TGCTCTCTTTCACACCTAGATTACCTAAAATAAGTTTCACTCAGAAAGCAACAAGAA 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200

QY 601 AATATCCATAAATGGTAAATGGTATGCTTGGGAAATGGTATATTATGGAGGCTCGGTT 660
Db |||||
201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyTrGlyGlySerGly 220
QY 661 AAACAACAGGCTCCATTCTTAACATATTCGCTCAAAATAAAACCCAGCTGGAGCCCTCCCAATG 720
Db |||||
221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProMet 240
QY 721 GCTATAGGACCAATACCGGCTTGAACGGGTCAAGACCCCAACCAACCAAGGACCCAGACCA 780
Db |||||
241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACAATACCTTGGATCAGACCCCACTAGTCTAAGACACCACTAAATGGG 840
Db |||||
261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
QY 841 GCATAACCTTTTAGCCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCAGACTCCAGAG 900
Db |||||
281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTTCTTGTGGCTATGCTTACGCTTCGGGCCCACTTACTACTAAGAAGGATGGCT 960
Db |||||
301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyTrGlyGlyMetAla 320
QY 961 AGAAGAGGGAATTCATGTGACAAAGAAACATAGAGACCAATGCACATGGGGATCCCAA 1020
Db |||||
321 ArgArgGlyLysPheAsnValThrLysGluHisArgPheGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTCGAAAGGCACCTGCATAGGAAAGGTTTCCCCCA 1080
Db |||||
341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACCAACCTTTGTAACACACTGAAGCTTTAATCAAACTCTGAGAGTCAATAT 1140
Db |||||
361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyTr 380
QY 1141 CTGGTACCTGTTATGACAGTGGTGGCAGTGTATCTGATTAACCCCTTGTGTTTCC 1200
Db |||||
381 LeuValProGlyTyTrAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTTAAACCAACTAAAGATTTTGGCATTATGGTCCCAATTTGTTCCCGAGTG 1260
Db |||||
401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTAATATCCGGAAGAAAGCAATCTTGATGATATGACTACAGAAATCATCGACAAAG 1320
Db |||||
421 TyTrTyTrProGlyLeuAlaIleLeuAspGluTyTrAspTyTrArgAsnHisArgGlnLys 440
QY 1321 AGAGAACCATATCTCTGACACTTCTGATGCTCGGACTTGGAGTGGCAGCAGGTGTA 1380
Db |||||
441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGGAACAGCTCGCTGTCTACGGGACCAACAGCAGCTAGAACAGGACTTAGTAAC 1440
Db |||||
461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATGTAAAGAGATCTCCAGCCCTAGAAAAATCTGTGAGTAACTGTGAG 1500
Db |||||
481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAAATCCCTAACCTCTTATCTGAAGTGTCTTACAGATAGAGAGGGGTAGATTATTA 1560
Db |||||
501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGATTTGTAGCTTTGAAGGAGGAATGCTGTTTTTATGTGGAT 1620
Db |||||
521 PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysPheTyTrValAsp 540
QY 1621 CATTCAGGGCCATCAGAGACTCCATGAACAGCTTTAGAGAAAGGTTGGAGAGGCTGCA 1680
Db |||||
541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluLeuGluLysArgArg 560
QY 1681 AGGGAAGGAAACTACTCAAGGGTGGTTTGAGGGATGGTTCAACAGACTCTCTTTGGTTG 1740

Db |||||
561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
QY 1741 GCTACCTACTTCTGCTTTAAACAGGACCCCTTAATAGTCTCTCTCTGTTACTACAGTT 1800
Db |||||
581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGSCCATGTATTATTAACNAAGTTAATTCCTTCATTAGAGAACCAAGTAAGTCCAGTCCAG 1860
Db |||||
601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGTGCTTACAGCAACAGTACCAAGCCCGCTCTAGCAGGGAAGCTGGCCGC 1914
Db |||||
621 IleMetValLeuArgGlnGlnTyTrGlnSerProSerSerArgGluAlaGlyArg 638
RESULT 9
ADY28041
ID ADY28041 standard; protein; 638 AA.
XX
AC ADY28041;
XX
DT 19-MAY-2005 (first entry)
XX
DE Porcine endogenous retrovirus subtype C ancestral env protein, SEQ ID 45.
XX
KW Xenotransplantation; vaccine; acquired immune deficiency syndrome;
KW anti-hiv; immune disorder; HIV-1 infection; infection; PERV infection;
KW diagnosis; graft versus host disease; immunosuppressive;
KW envelope protein.
XX
OS Porcine endogenous retrovirus; subtype C.
OS Synthetic.
XX
FN WO2005019411-A2.
XX
PD 03-MAR-2005.
XX
PF 19-MAY-2004; 2004WO-US015709.
XX
PR 19-MAY-2003; 2003US-00441949.
PA (AUCK-) AUCKLAND UNISERVICES LTD.
PA (UNIW) UNIV WASHINGTON.
PI Mullins JI, Rodrigo A, Ross HA;
XX
DR WPI; 2005-202625/21.
XX
PS New isolated ancestral viral nucleic acid sequence that is a determined
PT founder sequence of a highly diverse viral strain, subtype or group of an
PT endogenous retrovirus, useful in the prophylaxis of viral infection.
XX
Claim 10; SEQ ID NO 45; 278pp; English.
The invention relates to an isolated ancestral viral nucleic acid
sequence and its fragment, where the sequence is a determined founder
sequence of a highly diverse viral strain, subtype or group of an
endogenous retrovirus. Also included are an isolated ancestor protein (or
its fragment) from an endogenous retrovirus, an isolated expression
construct (comprising the following operably linked elements:- a
transcriptional promoter; a nucleic acid encoding an endogenous
retrovirus ancestor protein and a transcriptional terminator), a cultured
prokaryotic (or eukaryotic cell) transformed or transfected with the
expression construct, an isolated host cell comprising the expression
construct, a composition for inducing an immune response in a recipient
mammal (comprising a viral ancestor protein or its antigenic fragment,
where the protein is from a virus of a donor species), an isolated
antibody that binds specifically to an endogenous retrovirus ancestor
protein (and that binds specifically to a plurality of circulating
descendant endogenous retrovirus ancestor proteins), a method of
preparing an ancestral endogenous retroviral amino acid sequence, a
method for inducing an immune response to a donor virus in a transplant
recipient or a potential transplant recipient, a method of making a

CC vaccine, a method for detecting infection with an endogenous retrovirus
CC and a method for performing xenotransplantation in a subject. The
CC ancestral viral nucleic acid sequence is of Porcine Endogenous Retrovirus
CC (PERV) subtype A, B or C and has at least 70% identity with sequence of
CC any of the 18 nucleotide sequences fully defined in the specification.
CC The sequence may also be optimized for expression in a human host. The
CC nucleic acid sequence or its fragment is useful in the prophylaxis of
CC viral infection in transplantation that is heightened by the presence of
CC factors commonly associated with viral activation, e.g. immune
CC suppression, graft versus host disease, graft rejection, viral co-
CC infection, and cytotoxic therapies. The present sequence is an ancestral
CC PERV env protein.
XX
SQ

Sequence 638 AA;

Alignment Scores:

Pred. No.:	1.06e-289	Length:	638
Score:	3407.00	Matches:	635
Percent Similarity:	99.7%	Conservative:	1
Best Local Similarity:	99.5%	Mismatches:	2
Query Match:	98.2%	Indels:	0
DB:	9	Gaps:	0

US-10-723-552-3_COPY_5620_7533 (1-1914) x ADY28041 (1-638)

Qy	1	ATGCATCCCGAGTTAAACCGCGCCACCTCCCGATTCGGGTGGAAGCCGAAAGACTG	60
Db	1	MethisProThrLeuSerArgHisLeuProIleArgGlyGlyProLysArgLeu	20
Qy	61	AAATCCCTTAAGCTTCGCTCCATCGCGTGTCTTACTCTGTCATACCTCTCAG	120
Db	21	LysileProLeuSerPheAlaSerIleAlaIlePheLeuThrLeuSerIleThrSerGln	40
Qy	121	ACTAATGGTATGCCATAGGAGACAGCTGAACCTCCCAACACCTTATCTCTCACCTGG	180
Db	41	ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp	60
Qy	181	TTAATTACTGACTCCGCGACAGGTTAATATCAACCACTCAAGGGAGGCTCTTTA	240
Db	61	LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu	80
Qy	241	GGAACTGGTGGCTGATCTATACGTTTGGCTCGATCAGTATTCTCTGCTGACCTCA	300
Db	81	GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer	100
Qy	301	CCCCAGATATCTCCATGCTCAGGATTTATGTTGGCCAGGACCAACCAATAATGGA	360
Db	101	ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly	120
Qy	361	AAACATTGCGGAATCCAGAGATTTCTTTGTAAACAATGGAACCTGTGAACCTCTAAT	420
Db	121	LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn	140
Qy	421	GATGGATATTGGAATGCCACCTCTCAGCAGATAGGGTAAGTTTCTTATGTCAAC	480
Db	141	AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn	160
Qy	481	ACCTATACCACTCTGACACATTTAATTAACCTGACCTGGATTAGAACTGGAGCCCAAG	540
Db	161	ThrTyrThrSerSerGlyGlnPheAsnTrpLeuThrTrpIleArgThrGlySerProLys	180
Qy	541	TGCTCTCTCTCAGACCTAGATTACTAAAATAGTTTCACTGAGAAGGAAACCAAGAA	600
Db	181	CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu	200
Qy	601	AATATCTAAATGGGTAATGGTATCTCTTGGGAATGGTATATTATGAGGCTCGGGT	660
Db	201	AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGly	220
Qy	661	AAACACCGGCTCCATTCTAATCTTCGCTCAAAATAACCAACGCTGGAGCCTCCATG	720
Db	221	LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet	240

Qy	721	GCTATAGACCAAAATACGGTCTTACGGGTCAAAAGACCCCAACCCCAAGAGCAGGACCA	780
Db	241	AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro	260
Qy	781	TCCTCTAACATAACTCTGGATCAGCCCACTCAGTCTAACACACGACTAAATATGGG	840
Db	261	SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly	280
Qy	841	GCAAAACCTTTTACCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCACGACTCCAG	900
Db	281	AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu	300
Qy	901	GCTACCTCTCTTCTGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGAAATGGCT	960
Db	301	AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAla	320
Qy	961	AGAAAGGGGAAATCAATGTGACAAAGACATAGACCAATCAATCGGGATCCCAA	1020
Db	321	ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln	340
Qy	1021	AATAAGCTTACCTTACTGAGGTTTCTGGAAGGACCTGCATAGGAAAGGTTCCCCCA	1080
Db	341	AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro	360
Qy	1081	TCCCAACCAACCTTTGTAACCACTGAAGCCTTTAATCAAACTCTGAGAGTCAATAT	1140
Db	361	SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr	380
Qy	1141	CTGGTACCTGGTATGACAGGTGGTGGCATGTATTAATCTGATTAACCCCTTGTTTCC	1200
Db	381	LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer	400
Qy	1201	ACCTTGGTGTAAACCAAACTTAAAGATTTTGGATTATGGTCCAAATTTGTTCCCGAGT	1260
Db	401	ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal	420
Qy	1261	TATTACTATCCGAAAAAGCAATCTTGTATCAATATCACTACAGAAATCATCGACAAAG	1320
Db	421	TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys	440
Qy	1321	AGAAACCACTATCTCTGACACTGCTGTGATGCTCGGACTTGGAGTGGGAGGAGTGA	1380
Db	441	ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGlyVal	460
Qy	1381	GGAAACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
Db	461	GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn	480
Qy	1441	CTACATCGAATTTGTAACAGAAATCTCCAAAGCCTTAGAAAAATCTGTCACTGAG	1500
Db	481	LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu	500
Qy	1501	GAATCCCTAACCTCTTATCTGAAGTAGTCTCTACAGATAGAGAGGTTAGATTATTA	1560
Db	501	GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu	520
Qy	1561	TTTCTAAAACAAGAGGATTATGTAGCTCTGAAGGAGGAATGCTGTTTTTATGTGAT	1620
Db	521	PheLeuLysGlyGlyLeuCysValAlaLeuLysGlyGlyCysPheTyrValAsp	540
Qy	1621	CATTACGGGGCCATCAGAGACTCCATGAAACAAGCTTAGAGAAAGGTTGGAGAGCGTCA	1680
Db	541	HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg	560
Qy	1681	AGGAAAAGAAACTACTCAAGGGTGGTTGAGGGATGGTTCAACAGGTCTCTTTGGTGT	1740
Db	561	ArgGluLysGluThrThrGlnGlyTyrPheGluGlyTrpPheAsnArgSerProTrpLeu	580
Qy	1741	GCTTACCTCTCTCTTTTAAACAGGACCTTAATAGTCTCTCTCTCTCTCTCTCTCTCT	1800
Db	581	AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrVal	600
Qy	1801	GGGCCATGTATTATTAAACAAGTTAATTTGCTTCAATTAGAGAACGAATAAGTGCAGTCCAG	1860

Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGGTACTAGACAACAGTACCAAGCCGCTAGCAGGAAGCTGGCGC 1914
Db 621 IleMetValLeuArgGlnGlnTrpGlnSerProSerArgGluAlaGlyArg 638

RESULT 10
ADY28038
ID ADY28038 standard; protein; 638 AA.

XX AC ADY28038;
XX 19-MAY-2005 (first entry)
XX

DE Porcine endogenous retrovirus subtype C ancestral env protein, SEQ ID 42.
XX
XX Xenotransplantation; vaccine; acquired immune deficiency syndrome;
KW anti-hiv; immune disorder; HIV-1 infection; infection; PERV infection;
KW diagnosis; graft versus host disease; immunosuppressive;
KW envelope protein.

OS Porcine endogenous retrovirus; subtype C.
OS Synthetic.
XX WO2005019411-A2.

XX PD 03-MAR-2005.
XX PF 19-MAY-2004; 2004WO-US015709.

XX PR 19-MAY-2003; 2003US-00441949.
PA (AUCK-) AUCKLAND UNITSERVICES LTD.
PA (UNITW) UNIV WASHINGTON.

XX PI Mullins JI, Rodrigo A, Ross HA;
XX WPI; 2005-202625/21.
DR N-PSDB; ADY28037.

XX New isolated ancestral viral nucleic acid sequence that is a determined
PT founder sequence of a highly diverse viral strain, subtype or group of an
PT endogenous retrovirus, useful in the prophylaxis of viral infection.

PS Claim 10; SEQ ID NO 42; 278pp; English.

XX The invention relates to an isolated ancestral viral nucleic acid
CC sequence and its fragment, where the sequence is a determined founder
CC sequence of a highly diverse viral strain, subtype or group of an
CC endogenous retrovirus. Also included are an isolated ancestor protein (or
CC its fragment) from an endogenous retrovirus, an isolated expression
CC construct (comprising the following operably linked elements:- a
CC transcriptional promoter; a nucleic acid encoding an endogenous
CC retrovirus ancestor protein and a transcriptional terminator), a cultured
CC prokaryotic (or eukaryotic cell) transformed or transfected with the
CC expression construct, an isolated host cell comprising the expression
CC construct, a composition for inducing an immune response in a recipient
CC mammal (comprising a viral ancestor protein or its antigenic fragment,
CC where the protein is from a virus of a donor species), an isolated
CC antibody that binds specifically to an endogenous retrovirus ancestor
CC protein (and that binds specifically to a plurality of circulating
CC descendant endogenous retrovirus ancestor proteins), a method of
CC preparing an ancestral endogenous retroviral amino acid sequence, a
CC method for inducing an immune response to a donor virus in a transplant
CC recipient or a potential transplant recipient, a method of making a
CC vaccine, a method for detecting infection with an endogenous retrovirus
CC and a method for performing xenotransplantation in a subject. The
CC ancestral viral nucleic acid sequence is of Porcine Endogenous Retrovirus
CC (PERV) subtype A, B or C and has at least 70% identity with sequence of
CC any of the 18 nucleotide sequences fully defined in the specification.
CC The sequence may also be optimized for expression in a human host. The
CC nucleic acid sequence or its fragment is useful in the prophylaxis of

CC viral infection in transplantation that is heightened by the presence of
CC factors commonly associated with viral activation, e.g. immune
CC suppression, graft versus host disease, graft rejection, viral co-
CC infection, and cytotoxic therapies. The present sequence is an ancestral
XX PERV env protein.
SQ Sequence 638 AA;

Alignment Scores:
Pred. No.: 1.06e-289 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: 9 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x ADY28038 (1-638)

QY 1 ATGCATCCACGTTAAACCCGCGCCACCTCCCGGATTCGGGTGGAAAGCCGAAAGACTG 60
Db 1 MethisProThrLeuSerArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAAGCTTCGCTCCATCGCGGTTCCTTACTCTGTCAATTAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 40
QY 121 ACTAATGGTATGCGCATAGGACAGCTGACCTGACCTCAATCAACCTTATCTCACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGCGCACAGGTATTAAATCAACCACTCAAGGCGGCTCTTTA 240
Db 61 LeulleThrAspSerGlyThrGlyIleAsnIleAsnIleAsnIleAsnIleAsnIleAsnIle 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTATTTCCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCTCCATGCTCAGGATTTTATGTTTGGCCGAGCACCACCAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGCGAAATCCAGAGATTCTTTTGTAAACAATCGAACTGTGAACCTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTCGAAATGCCAACCTCTCAGCAGATAGGTAAGTTTTCTTATGTCTAAC 480
Db 141 AspGlyTyTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyValAsn 160
QY 481 ACCTATACCAGCTCTGGACATTTAATTAACCTGACCTGGATTAGAACTGGAGCCCAAG 540
Db 161 ThrTyThrSerSerGlyGlnPheAsnTyLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCTCCAGACCTAGATTACTCTAAAAATAAGTTTCACTGAGAAAGGAAAAACAAGA 600
Db 181 CysSerProSerAspLeuAspTyLeuLysIleSerPheThrGluLysGlyGlnGlu 200
QY 601 AATATCTTAAATGGTAAATGGTATGTTCTTGGGAATGGTATATATATGAGGCTCGGGT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyTrGlyGlySerGly 220
QY 661 AAACACAGGCTCCATCTTAACCTATTCGCTCAAAATAAACCCAGCTGGAGCCTCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGACCAATAATACGTTCTTCGCGGTCAAAGACCCCAACCCAGGACCCAGACCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACATAACTTCTGGATCAGACCCCACTAGTCTAACAGCACCACTAAATATGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280

Qy	841	GC	AAAACTTTT	AGCCTCATCCAGGGAGC	TTTCAAGCTCTTA	ACTCCACGACTCCAGAG	900
Db	281	Ala	lysLeuPheSerLeu	IleGlnGlnGlnAlaPheGlnAlaLeu	AsnSerThrThrProGlu	300	
Qy	901	GCT	ACCTCTTCT	TGCTGGCTATGCTTAGCTTCGGGGCCAC	CTTACTATGAAGGAATGGCT	960	
Db	301	Ala	ThrSerSerCys	TrpLeuCysLeuAlaLeuGlyPro	ProTyrTyrGluGlyMetAla	320	
Qy	961	AGA	AGAGGGAAATT	CAATGTGACAAAAGAAACATAGACCA	CAATCGGAGTCCCAA	1020	
Db	321	Arg	ArgGlyLysPhe	AsnValThrLysGluHisArgAspGlnCys	ThrTrpGlySerGln	340	
Qy	1021	AAT	AAAGCTTAC	CTTACTGAGGTTTCTGGAAAGGCACCTGC	ATAGAAAGGTTCCCCCA	1080	
Db	341	Asn	LysLeuThrLeu	ThrGluValSerGlyLysGlyThrCys	IleGlyLysValProPro	360	
Qy	1081	TCC	CACCAACACCT	TTGTAAACCAACACTGAAGCCCTTTAATCA	AAACCTCTGAGAGTCAATAT	1140	
Db	361	Ser	HisGlnHisLeu	CysAsnHisThrGluAlaPheAsnGln	ThrSerGluSerGlnTyr	380	
Qy	1141	CTG	GTACTCGTT	TANGACAGGTGGGCATGTAATACTGGA	TAAACCCCTTGTTTTC	1200	
Db	381	Leu	ValProGlyTyr	AspArgTrpTrpAlaCysAsnThrGly	LeuThrProCysValSer	400	
Qy	1201	ACC	TGTGTTTTT	TACCAAACTAAAGATTTTTCATATATG	TGCCAAATGTTTCCCCGATG	1260	
Db	401	Thr	LeuValPheAsn	GlnThrLysAspPheCysIleMetValGln	IleValProArgVal	420	
Qy	1261	TAT	TACTATCCC	GAAGCAACTCTTGTGATGAATATGACTAC	AGAAATCATCGACAAAAG	1320	
Db	421	Tyr	TyrTyrPro	GluLysAlaIleLeuAspGluTyrAspTyr	ArgAsnHisArgGlnLys	440	
Qy	1321	AGA	AAACCAATCT	CTGACACTTCTGTGTGATGCTCGGACTTG	GAGTGGCAGCAGGTGTA	1380	
Db	441	Arg	GluProIleSerLeu	ThrLeuAlaValMetLeuGlyLeuGlyVal	AlaIleGlyVal	460	
Qy	1381	GGA	ACGAAACAG	CTGCCCTGCTCACGGGACACAGCAGCT	TAGAAACAGACTTACTNAC	1440	
Db	461	Gly	ThrGlyThrAla	AlaLeuValThrGlyProGlnGlnLeu	GluThrGlyLeuSerAsn	480	
Qy	1441	CTA	CTCGAATTGT	TAAAGAGAATCTCCAAGCCCTAGAAAAT	CTGTCACTAACCTGGAG	1500	
Db	481	Leu	HisArgIleVal	ThrGluAspLeuGlnAlaLeuGluLys	SerValSerAsnLeuGlu	500	
Qy	1501	GAAT	CCCTAACCT	CTTATCTGAAGTAGTCTCTACAGATAG	AGAGGGTTAGATTATTA	1560	
Db	501	Glu	SerLeuThrSerLeu	SerGluValValLeuGlnAsnArgArgGly	LeuAspLeuLeu	520	
Qy	1561	TTT	CTAAAAGA	GAGGATTATGTGTAGCCTCGAAGGAA	AGTCGTCTTTATGTGGAT	1620	
Db	521	Phe	LeuLysGluGly	LysCysValAlaLeuLysGlnGluCys	PheTyrValAsp	540	
Qy	1621	CAT	TCAAGGGCC	ATCAGAGACTCCATGAACAAGCTTAGA	AAAAGTTGGAGAGCGCTCGA	1680	
Db	541	His	SerGlyAlaIle	ArgAspSerMetAsnLysLeuArgGlu	ArgLeuGluLysArgAG	560	
Qy	1681	AGG	AAAGAAACT	ACTCAAGGTGTTTGAGGATGTTTCAAC	AGGTCTCTTTGGTTG	1740	
Db	561	Arg	GluLysGluThr	ThrGlnGlyTyrPheGluGlyTrpPhe	AsnArgSerProTyrLeu	580	
Qy	1741	GCT	ACCTACTT	CTCTGCTTTAACAGACCTCTTAATAGT	CTCTCTCTGTTTACTCACAGTT	1800	
Db	581	Ala	ThrLeuLeuSer	AlaLeuThrGlyProLeuIleValLeu	LeuLeuLeuLeuThrVal	600	
Qy	1801	GGG	CCATGTATT	ATAAGTTAATTTGCCCTTCATTAGAGA	ACGAATAAGTCAGTCCAG	1860	
Db	601	Gly	ProCysIleIle	AsnLysLeuIleAlaPheIleArgGlu	ArgIleSerAlaValGln	620	
Qy	1861	ATC	ATGGTACT	TAGACACAGTACCAAGCCCTCTAG	CAGGGAAGCTGCCGC	1914	
Db	621	Ile	MetValLeuArg	GlnGlnTyrGlnSerProSerArgGlu	AlaGlyArg	638	

RESULT 11	
AAB70634	standard; protein; 638 AA.
ID	AAB70634
XX	
AC	AAB70634;
XX	
DT	15-MAY-2001 (first entry)
XX	
DE	Porcine endogenous retrovirus envelope protein TSUKUBA.
XX	
KW	Porcine endogenous retrovirus; PoERV; gag; env; detection; antigenic;
KW	immunogenic; antiviral; vaccine; antiserum; viral infection.
XX	
OS	Porcine endogenous retrovirus.
XX	
FN	WO200112816-A2.
XX	
PD	22-FEB-2001.
XX	
PF	16-AUG-2000; 2000WO-GB003159.
XX	
PR	18-AUG-1999; 99GB-00019604.
XX	
PA	(QUIP-) QUIP TECHNOLOGY LTD.
XX	
FI	Galbraith D, Kelly H, Smith K;
XX	
DR	WPI; 2001-211223/21.
XX	
PT	New porcine endogenous retrovirus polypeptide fragment with retroviral
PT	specific antigenic or immunogenic activity, for detection of retroviral
PT	antibodies in a sample and in therapy or diagnosis.
XX	
PS	Claim 7; Fig 2; 43pp; English.
XX	
CC	The present invention describes a porcine endogenous retrovirus (PoERV)
CC	polypeptide fragment (I) which has PoERV specific antigenic or
CC	immunogenic activity. Also described are: (1) an antiserum (II) specific
CC	to (I); (2) a PoERV specific antibody (III) or its fragment raised
CC	against (I); (3) use of (I) in detecting (III) in a sample; (4) use of
CC	(III) in detecting PoERV in a sample; (5) use of (I) or (III) in therapy
CC	or diagnosis; and (6) an assay kit (IV) comprising (I) or (III) for use
CC	in detection of PoERV in a sample. (I) has antiviral activity, and can be
CC	used in vaccine production. (I) is useful in the detection of PoERV
CC	antibodies (III) in a sample and (III) is useful in the detection of
CC	PoERV in a sample. (I) and (III) are useful in therapy or diagnosis.
CC	Antiserum (II) to (I) and kits (IV) comprising (I) or (III) are useful in
CC	detection of specific PoERV types. Viral infection can be monitored and
CC	by using (I) or (III). The present sequence represents a specifically
CC	claimed PoERV envelope (env) protein sequence from the present invention
XX	
SQ	Sequence 638 AA;
Alignment Scores:	
Pred. No.:	2.92e-289 Length: 638
Score:	3402.00 Matches: 634
Percent Similarity:	99.7% Conservative: 2
Best Local Similarity:	99.4% Mismatches: 2
Query Match:	98.1% Indels: 0
DB:	4 Gaps: 0
US-10-723-552-3_COPY_5620_7533 (1-1914) x AAB70634 (1-638)	
Qy	1 ATGATCCCAAGTTAAACCGCGCACCTCCCGATTCCGGGTGGAAACCGAAGACTG 60
	:::
Db	1 MethisProthrLeuSerArgHisLeuProIleArgGlyGlyLeuProLysArgLeu 20
Qy	61 AAAATCCCTTAAGCTTCGCCTCCATCGCGGGTTCCTTACTCTGTCATAAACCTCTCAG 120
Db	21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
Qy	121 ACTAATGGTATGGCATAGGAGACAGCTGAACCTCCCAATAACCCCTTATCTCTCACCTGG 180

Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCGGCACAGGTATTAAATATCAACACTCAAGGGAGGCTCCCTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTTGATCTACCTTTGCTCCTCAGATCAGTATTATTCCTAGTCTCAGCTCA 300
Db 81 GlyThrTrpTrpProAspLeuThrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCCTCCATCTCAGCGATTTTATGTTTCCCGAGGACCAACCAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGGAATCCAGAGATTTCTTTTGTAAACAATGGAAGTGTGAACCTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGGATATTGGAATGGCCAACTCTCAGCAGGATAGGTAAGTTTCTTATGTCAAC 480
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACCACTCTGGCAATTTAATTACCTGACCTGGATTAGAACTGGAAGCCCCAAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCTCAGACTGAGATTACCTAAATAAGTTTCTACCTGAGAAAGGAAACAAGAA 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTTAAATGGTAAATGGTATGCTTGGGAATGGTATATTATGGAGGCTCGGGT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyLysSerGly 220
QY 661 AAACAACCAAGGCTCCATCTCACTATTCGCTCAAAATAAACCACTGAGGCTCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProMet 240
QY 721 GCTATAGGACCAATACGGTCTTGAGGGTCAAGACCCCAACCCCAAGGACCGAGCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACTAACTCTCGGATCAGACCCCACTGAGTCTTAACAGCACCACTAAATGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerSerThrThrLysMetGly 280
QY 841 GCAAACTTTTTCAGCTCATCCAGGGAGCTTTTCAAGCTCTTAACTCCACGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACTCTTCTTGTGGTATGCTTAGCTTCGGGCCCACTTACTACTATGAGGAATGGCT 960
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAla 320
QY 961 AGAAGCGGAATTCAAATGTGAAAGAAACAACATAGAGACCAATGACATGGGATCCCAA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTGGAAGGCACCTGCATAGGAAGGTTCCCCCA 1080
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACCAACACTTTGTAAACCACTAGAGCTTTAATCAAACTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGGTACCTGGTTATGACAGGTGGGCACTCATCTATGATTAAACCCCTTGTCTTCC 1200
Db 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTAAACCAACTAAAGATTTTTCATTATGGTCCAAATGTTCCCGGAGTG 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420

QY 1261 TATTACTATCCGAAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAAG 1320
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAAGACCCATATCTCTGACACTTGTCTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGA 1380
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGAAACAGCTGCCCTGCTCAGGGACCAACAGCAGCTAGAAAACAGGACTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTGTAAACAGAAAGATCTCAAGCCCTAGAAAAATCTGTCAGTAACCTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAAATCCCTAACTCTCTTATCTGAAGTAGTCTCTCAGAAATAGAAGAGGTTAGATTTATTA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAAGAGGAGGATTTATGTAGCTTGTAGCCTTGAAGGAGGAATCTGTTTTTATGTGGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
QY 1621 CATTCAGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAAAGGTTGGAGAAAGCGTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGGAAGAGGAAACTACTCAAGGGTGGTTTGGAGGATGGTTCAACAGGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
QY 1741 GCTACCTTACTTTCTGCTTTAAACAGGACCTTATAGTCTCTCTCTCTCTCTCTCTCTCT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATCTATTATTAAACAGTAAATTCCTTATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGTGTACTTAGACAACAGTACCAAGCCCGTCTAGCAGGGAAGCTGGCCGC 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 638
RESULT 12
ADSS73449
ID ADSS73449 standard; protein; 666 AA.
XX AC ADSS73449;
XX DT 16-DEC-2004 (first entry)
XX DE Swine retroviral (Taukuba-1) env protein.
XX KW Swine retroviral protein; immunosuppressive; gene therapy; env protein.
XX OS Porcine endogenous retrovirus.
XX FH Key Location/Qualifiers
FT Misc-difference 13 /note= "Encoded by TGA"
FT FT Misc-difference 19 /note= "Encoded by TGA"
XX PN US2004185435-A1.
XX PD 23-SEP-2004.
XX PF 26-NOV-2003; 2003US-00723552.
XX PR 14-DEC-1995; 95US-00572645.
PR 13-DEC-1996; 96US-00766528.

201	LeuAspTyrLeuLysIleSerPheThrGluLysGlyLeuAsnIleLeuLysTrp	220
616	GTAAATGGTATGCTTGGGGAATGGTATATTATGGAGGCTCGGGTAAACAAACACGAGCTCC	675
221	ValAsnGlyMetSerTyrGlyMetValTyrGlySerGlyLysGlnProGlySer	240
676	ATTCTAACTATTGGCTCAAAATAAAC---CAGCTGGAGCCTCCAAATGGCTATAGGACCA	732
241	IleLeuThrIleArgLeuLysIleAsnThrGlnLeuGluProProMetAlaIleGlyPro	260
733	AATACGGTCTGACGGGTCAAGACCCCAACCCCAAGGACCA-----	774
261	AsnThrValLeuThrGlyLysArgProThrGlnGlyProProHisAsnLeuProVal	280
775	-----GGACCATCTCT-----	801
281	ProGlnGlyProSerProAsnProAspIleThrGlnSerAspTyrAsnIleThrSerGly	300
802	TCAGACCCCACT-----GAGTCTTAACAGACGACTAAATAATGGGGCAAAA	840
301	SerAspProThrAsnThrProArgAsnGluSerAsnSerThrThrLysMetGlyAlaLys	320
847	CTTTTAAAGCCTCATCCAGGAGCTTTCAAGCTCTTAAGCTCCAGACTCCAGAGGCTACC	906
321	LeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThr	340
907	TCTTCTTGTGGCTATGCTTACGTTCTGGGGCCACCTTACTATGAAGGAATGGCTAGAAGA	966
341	SerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAlaArgArg	360
967	GGGAAATCTAATGTGCAAAAAGAACATAGAGACCAATGCACATGGGGATCCCAAAATAAG	1026
361	GlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTyrGlySerGlnAsnLys	380
1027	CTTACCTCTTACTGAGGTTCTCGAAAGGACCTGCATAGAAAGGTTCCCCCATCCAC	1086
381	LeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHis	400
1087	CAACACCTTTGTAAACACACTGAAGCCTTTAATCAAACTCTGAGAGTCAATATCTGTA	1146
401	GlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyrLeuVal	420
1147	CCTGGTTATGACAGGTGGTGGCATGTAATCTGGATTAACCCCTTGTTGTTCCACCTTG	1206
421	ProGlyTyrAspArgTyrTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeu	440
1207	GTTTTAAACCAACTAAGATTGTTGCAATTATGCTCCAAATGTTCCCGAGGTATTAC	1266
441	ValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrTyr	460
1267	TATCCCCAAAAAGCAATCCTTGATGAATATGACTACAGAATATCATCAGAAAAGAGAGAA	1326
461	TyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLysArgGlu	480
1327	CCCATACTCTGACACTTGCTGTGATGCTCGACTTGGAGTGGCAGCAGGTGTAGGAACA	1386
481	ProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThr	500
1387	GGAAACAGCTGCCCTGGTACGGGACCAACAGCAGCTAGAAAACAGACTTAGTAACTCAT	1446
501	GlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHis	520
1447	CGAATTGTAACAGAAGATCTCCAAGCCCTAGAAAAATCTGTCAGTAACCTGGAGGAATCC	1506
521	ArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSer	540
1507	CTAACCTCTCTTATCTGAAGTAGTCCTACAGAAATAGAGAGGGTTAGATTATTATTTCTA	1566
541	LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu	560
1567	AAGAAGAGGAGGATTATGTGTAGCTTGAAGAGGAATCTGTTTTTATGTGGATCATCA	1626
561	LysGluGlyLysValAlaLeuLysGluGluCysCysPheTyrValAspHisSer	580

QY	1627	GGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAAAGGTTGGAGAGAGCCTCGAAGGAA	1686
Db	581	GlyAlaIleArgAspSerMetAsnLysLeuArgLeuGluArgLeuGluLysArgArgGlu	600
QY	1687	AAGGAAACTACTCAAGGTCGTTTGAGGGATGTTCAACAGGTCTCTTCGTGGCTACC	1746
Db	601	LysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeuAlaThr	620
QY	1747	CTACTTCTCTCTTTAAACAGGACCCCTTAATAGTCTCTCTCTCTTACTTCACAGTTGGGCCA	1806
Db	621	LeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrValGlyPro	640
QY	1807	TGTATTATTAAACAAGTTAATTCCTTCATTAGAGACGAATAAGTCAGTCCAGATCATG	1866
Db	641	CysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMet	660
QY	1867	GTACTTAGACAACAGTACCAAAAGCCGTCCTAGCAGGGAAGCTGCGCGC	1914
Db	661	ValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg	676
RESULT	14		
ADY	28034		
ID	ADY28034	standard; protein; 678 AA.	
XX			
AC	ADY28034;		
XX			
DT	19-MAY-2005	(first entry)	
XX			
DE	Porcine endogenous retrovirus subtype C ancestral env protein, SEQ ID 38.		
XX			
KW	Xenotransplantation; vaccine; acquired immune deficiency syndrome;		
KW	anti-hiv; immune disorder; HIV-1 infection; infection; PERV infection;		
KW	diagnosis; graft versus host disease; immunosuppressive;		
KW	envelope protein.		
XX			
OS	Porcine endogenous retrovirus; subtype C.		
OS	Synthetic.		
XX			
PN	WO2005019411-A2.		
XX			
PD	03-MAR-2005.		
XX			
XX	19-MAY-2004; 2004WO-US015709.		
PF			
XX			
PR	19-MAY-2003; 2003US-00441949.		
XX			
PA	(AUCK-) AUCKLAND UNISERVICES LTD.		
PA	(UNIW) UNIV WASHINGTON.		
XX			
PI	Mullins JI, Rodrigo A, Ross HA;		
XX			
DR	WPI; 2005-202625/21.		
DR	N-PSDB; ADY28033.		
XX			
PT	New isolated ancestral viral nucleic acid sequence that is a determined		
PT	founder sequence of a highly diverse viral strain, subtype or group of an		
PT	endogenous retrovirus, useful in the prophylaxis of viral infection.		
XX			
PS	Claim 10; SEQ ID NO 38; 278pp; English.		
XX			
CC	The invention relates to an isolated ancestral viral nucleic acid		
CC	sequence and its fragment, where the sequence is a determined founder		
CC	sequence of a highly diverse viral strain, subtype or group of an		
CC	endogenous retrovirus. Also included are an isolated ancestor protein (or		
CC	its fragment) from an endogenous retrovirus, an isolated expression		
CC	construct (comprising the following operably linked elements:- a		
CC	transcriptional promoter; a nucleic acid encoding an endogenous		
CC	retrovirus ancestor protein and a transcriptional terminator), a cultured		
CC	prokaryotic (or eukaryotic cell) transformed or transfected with the		
CC	expression construct, an isolated host cell comprising the expression		
CC	construct, a composition for inducing an immune response in a recipient		
CC	mammal (comprising a viral ancestor protein or its antigenic fragment.		

CC where the protein is from a virus of a donor species), an isolated
 CC antibody that binds specifically to an endogenous retrovirus ancestor
 CC protein (and that binds specifically to a plurality of circulating
 CC descendant endogenous retrovirus ancestor proteins), a method of
 CC preparing an ancestral endogenous retroviral amino acid sequence, a
 CC method for inducing an immune response to a donor virus in a transplant
 CC recipient or a potential transplant recipient, a method of making a
 CC vaccine, a method for detecting infection with an endogenous retrovirus
 CC and a method for performing xenotransplantation in a subject. The
 CC ancestral viral nucleic acid sequence is of Porcine Endogenous Retrovirus
 CC (PERV) subtype A, B or C and has at least 70% identity with sequence of
 CC any of the 18 nucleotide sequences fully defined in the specification.
 CC The sequence may also be optimized for expression in a human host. The
 CC nucleic acid sequence or its fragment is useful in the prophylaxis of
 CC viral infection in transplantation that is heightened by the presence of
 CC factors commonly associated with viral activation, e.g. immune
 CC suppression, graft versus host disease, graft rejection, viral co-
 CC infection, and cytotoxic therapies. The present sequence is an ancestral
 CC PERV env protein.

XX Sequence 678 AA;

Alignment Scores:

Pred. No.: 7,06e-282 Length: 678
 Score: 3318.00 Matches: 634
 Percent Similarity: 93.9% Conservative: 1
 Best Local Similarity: 93.8% Mismatches: 3
 Query Match: 95.7% Indels: 38
 DB: 9 Gaps: 7

US-10-723-552-3_COPY_5620_7533 (1-1914) x ADV28034 (1-678)

QY 1 ATGCATCCACGCTTAACCGCGGCCACCTCCGAGTTCGGGTGGAAACCCGAAAGACTG 60
 DB 1 MethisProThrLeuSerArgHisLeuProIleArgGlyGlyLeuProLysArgLeu 20
 QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCTTACTCTCAATAACCTCTCAG 120
 DB 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
 QY 121 ACTAATGGTATGGCATAGGACAGCCTGAACCTCCCATAAACCTTATCTCTCACCTGG 180
 DB 41 ThrAenGlyMetArgIleGlyAspSerLeuAenSerHisLysProLeuSerLeuThrTrp 60
 QY 181 TTAATTACTGACTCGGCACAGGTATTAATATCAACACTCAAGGGAGGCTCCTTTA 240
 DB 61 LeuileThrAspSerGlyThrGlyIleAenIleAenAenThrGlnGlyAlaProLeu 80
 QY 241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTATTCCTAGTCTG----- 294
 DB 81 GlyThrTrpTrpProAspLeuTyrrValCysLeuArgSerValIleProSerLeuAenAsp 100
 QY 295 ---ACCTCACCCCCAGATATCTCCATGCTCACGGATTTTATGTTGGCCAGGACCACCA 351
 DB 101 GlnThrSerProProAspIleLeuHisAlaHisGlyPheTyrrValCysProGlyProPro 120
 QY 352 AATAATGGAACACATGCGGAAATCCAGAGATTTCTTTGTAAACAATGGAAGTGTGTA 411
 DB 121 AenAenGlyLysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysVal 140
 QY 412 ACCTCTAATGATGGATATTGGAAATGGCAACCTCTCAGCAGGATAGGTTAAGTTTCT 471
 DB 141 ThrSerAenAspGlyTyrrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSer 160
 QY 472 TATGTCAACACCTATACCAGCTCTCGACAATTAATTAC-----CTGACC 516
 DB 161 TyrValAenThrTyrrThrSerSerGlyGlnPheAenTyrrGlyHisGlyArgTrpLeuThr 180
 QY 517 TGG-----ATTAGAAGTGAAGCCCAAGTCTCTCCTTCAGAC 555
 DB 181 TrpGlnGlnArgValGlnLysAspIleArgThrGlySerProLysCysSerProSerAsp 200
 QY 556 CTAGATTACCTAAAAATAAGTTTCACTGAGAAAGGAAAAACAAGAAATATCTCAAAATGG 615

DB 201 LeuAspTyrrLeuLysIleSerPheThrGluLysGlyLysGlnGluAenIleLeuLysTrp 220
 QY 616 GTAATATGGTATCTTCTGGGGAATCGTATATATATGAGGCTCGGTAAACACCAAGGCTCC 675
 DB 221 ValAsnGlyMetSerTrpGlyMetValTyrrGlySerGlyLysGlnProGlySer 240
 QY 676 ATTCTAACTATTCCCTCAAAATAAAC---CAGCTGGAGCCTCCAATGGCTATAGGACCA 732
 DB 241 IleLeuThrIleArgLeuLysIleAenThrGlnLeuGluProProMetAlaIleGlyPro 260
 QY 733 AATACGGTCTTTCAGCGGTCAAAAGACCCCAACCAAGGACCA----- 774
 DB 261 AenThrValLeuThrGlyGlnArgProProThrGlnGlyProProHisAenLeuProVal 280
 QY 775 -----GGACCATCTCT----- 801
 DB 281 ProGlnGlyProSerProAenProAspIleThrGlnSerAspTyrrAenIleThrSerGly 300
 QY 802 TCAGACCCCACT-----GAGTCTAACAGCAGCTAAAAATGGGGCAAAA 846
 DB 301 SerAspProThrAenThrProArgAsnGluSerAenSerThrThrLysMetGlyAlaLys 320
 QY 847 CTTTTTAGCCTCATCCAGGGAGCTTTTCAAGCTCTTAACTCCACGACTCCAGAGGCTACC 906
 DB 321 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAenSerThrThrProGluAlaThr 340
 QY 907 TCTTCTTGTGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGNAATGGCTAGAGA 966
 DB 341 SerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrrTyrrGluGlyMetAlaArg 360
 QY 967 GGGAAATCAATGACAAAGAACATAGAGACCAATGCACATGGGATCCCAAAATAAG 1026
 DB 361 GlyLysPheAenValThrLysGluHisArgaspGlnCysThrTrpGlySerGlnAsnLys 380
 QY 1027 CTTACCCCTTACTGAGGTTTCTGGAAAAGGCACTGTCATAGGAAAGGTTTCCCCCATCCCAC 1086
 DB 381 LeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHis 400
 QY 1087 CAACACCTTTGTACCACACACTGAAGCCTTTAATCAACCTCTGAGAGTCAATATGTGTA 1146
 DB 401 GlnHisLeuCysAenHisThrGluAlaPheAenGlnThrSerGluSerGlnTyrrLeuVal 420
 QY 1147 CTTGTTATGACAGCTGGTGGGCATGTAATACTGGATTAACCCCTTGTTGTTTCCACCTTG 1206
 DB 421 ProGlyTyrrAspArgTrpTrpAlaCysAenThrGlyLeuThrProCysValSerThrLeu 440
 QY 1207 GTTTTAAACCAAACTAAAGATTTTTCATTATGTCCTCAAAATGTTTCCCGAGTGTATTAC 1266
 DB 441 ValPheAenGlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrr 460
 QY 1267 TATCCGMAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAGAGACA 1326
 DB 461 TyrProGluLysAlaIleLeuAspGluTyrrAspTyrrArgAenHisArgGlnLysArgGlu 480
 QY 1327 CCCATATCTCGACACTTGTCTGATGCTCGGACTTGGAGTGGCAGCAGGTGTAGGAACA 1386
 DB 481 ProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGlyValGlyThr 500
 QY 1387 GGAACAGCTGCCCTCGGTCAAGGACCCACAGCAGCTAGAAAACAGGACTTTAGTAACCTCAT 1446
 DB 501 GlyThrAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAenLeuHis 520
 QY 1447 CGAATTGTAAACAGAGATCTCAAGCCCTAGAAAAATCTGTCAGTAACTCGAGGAATCC 1506
 DB 521 ArgIleValThrGluAspLeuGlnAlaLeuGlyLysSerValSerAenLeuGluGluSer 540
 QY 1507 CTACCTCTCTTATCTGAGTAGTCTCAGAGATAGAAGAGGCTTAGATTTATTATTCTTA 1566
 DB 541 LeuThrSerLeuSerGluValValLeuGlnAenArgGlyLeuAspLeuPheLeu 560
 QY 1567 AAAGAAGGAGGATTTATGTGTAGCCCTTGAAGGAGGAATGCTGTTTTTATGTGGATCATTTCA 1626


```
Db      321  TriLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAlaAArgArgGlyLysPhe 340
QY      976  AATGTGACAAAGAACATAGAGACCAATGACATGGGATCCCAAAATAAGCTTACCCCTT 1035
Db      341  AsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLysLeuThrLeu 360
QY     1036  ACTGAGGTTTCTGGAAAAGGCACCTGCATAGGAAGGTTCCCCCATCCACCAACACCTT 1095
Db     361  ThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu 380
QY     1096  TGTAAACCACTGAAGCCCTTAAATCAAAACCTCTGAGAGTCAATATCTGCTACCTGGTTAT 1155
Db     381  CysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyrLeuValProGlyTyr 400
QY     1156  GACAGTGGTGGGCATGTAATACTAGATTAAACCCCTTGTGTTCCACCTTGGTTTTTAAC 1215
Db     401  AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn 420
QY     1216  CAAACTAAGATTTTGGCATTATGTCCTCAAAATTTGTTCCCGAGTGTATTACTATCCGAA 1275
Db     421  GlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrTyrProGlu 440
QY     1276  AAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAGAGAGAACCCATATCT 1335
Db     441  LysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLysArgGluProIleSer 460
QY     1336  CTGACACTTGTCTGTGATCTCGACTTGGAGTGGCAGAGGTGTAGGAACAGGAACAGCT 1395
Db     461  LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThrAla 480
QY     1396  GCCCTGTCACGGGACCAACGACGCTAGAAACAGGACTTAGTAACTACATCGAATTGTA 1455
Db     481  AlaLeuValThrGlyProGlnLeuGluThrGlyLeuSerAsnLeuHisArgIleVal 500
QY     1456  ACAGAAAGATCTCCAAGCCCTAGAAAAATCTGTCACTAACTGGAGGAATCCCTAACCTCC 1515
Db     501  ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSerLeuThrSer 520
QY     1516  TTATCTGAAGTAGTCCTCAGAAATAGAAGAGGTTAGATTATTATTCTAAAAAGAGGA 1575
Db     521  LeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuPheLeuLysGluGly 540
QY     1576  GGATTATGTGTAGCTTTGAAGGAGCAATGCTGTTTTTTATGTGGATCATTACAGGGCCATC 1635
Db     541  GlyLeuCysValAlaLeuLysGluCysCysPheTyrValAspHisSerGlyAlaIle 560
QY     1636  AGAGACTCCATGAACAAGCTTAGAGAAAGGTTGAGAGACGTCGAAAGGGAAGGAAACT 1695
Db     561  ArgAspSerMetSerLysLeuArgGluArgLeuGluLysArgArgGluLysGluThr 580
QY     1696  ACTCAAGGGTGGTTGAGGGATGGTTCAACAGGTCTCTTTGGTTGGCTACCCCTACTTTCT 1755
Db     581  ThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeuAlaThrLeuLeuSer 600
QY     1756  GCTTTTAAAGGACCCCTTAAATAGTCCTCTCTGTTACTCACAGTTGGGCCATGTATTATT 1815
Db     601  AlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrValGlyProCysIleIle 620
QY     1816  AACAAAGTTAATTCCTTCATTAGAGAACGAATAAGTCAGTCCAGATCATGTGTACTTAGA 1875
Db     621  AsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMetValLeuArg 640
QY     1876  CAACAGTACCAAGCCCTCTAGCAGGGAAGCTGCCCGC 1914
Db     641  GlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 653
```

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:07:51 ; Search time 48.1266 Seconds
(without alignments)
3323.427 Million cell updates/sec

Title: US-10-723-552-3_COPY_5620_7533

Perfect score: 3468

Sequence: 1 ATGCATCCAGCTTAACCG.....CTAGCAGGAGAGCTGGCGCGC 1914

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=x1p
-Q=/abes/ABSSWEB/spool/US10723552/runat_14022006_125148_13205/app_query.fasta_1
-DB=Published Applications AA Main -QPMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum2
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abes02p
-USER=US10723552 @CNG 1.1 602 @runat_14022006_125148_13205 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3420	98.6	638	3	US-09-851-859A-6
2	3407	98.2	638	4	US-10-441-949-32
3	3407	98.2	638	4	US-10-441-949-34
4	3407	98.2	638	4	US-10-441-949-40
5	3407	98.2	638	4	US-10-441-949-42
6	3407	98.2	638	4	US-10-441-949-45
7	3318	95.7	678	4	US-10-441-949-36
8	3318	95.7	678	4	US-10-441-949-38
9	3147.5	90.8	653	4	US-10-029-656-2
10	3135.5	90.4	653	4	US-10-029-656-4
11	2879.5	83.0	661	4	US-10-441-949-43

12	2874.5	82.9	660	4	US-10-441-949-16
13	2871.5	82.8	660	4	US-10-441-949-8
14	2866.5	82.7	660	3	US-09-851-859A-5
15	2851	82.2	678	4	US-10-441-949-12
16	2849.5	82.2	660	4	US-10-441-949-10
17	2849.5	82.2	660	4	US-10-441-949-18
18	2835	81.1	678	4	US-10-441-949-14
19	2465	70.8	678	4	US-10-441-949-26
20	2455	70.8	678	4	US-10-441-949-24
21	2432	70.1	657	4	US-10-441-949-22
22	2432	70.1	657	4	US-10-441-949-30
23	2432	70.1	658	4	US-10-441-949-44
24	2422	69.8	657	3	US-09-851-859A-4
25	2422	69.8	657	4	US-10-441-949-20
26	2422	69.8	657	4	US-10-441-949-28
27	2353	67.8	656	3	US-09-851-859A-3
28	1520	43.8	676	4	US-10-677-558-3
29	1489.5	42.9	667	3	US-09-070-630-13
30	1489.5	42.9	667	5	US-10-732-438-13
31	1347	38.8	672	3	US-09-954-983-4
32	1346	38.8	675	3	US-09-954-983-6
33	1346	38.8	675	3	US-09-954-983-8
34	1345	38.8	672	3	US-09-954-983-5
35	1345	38.8	676	3	US-09-954-983-11
36	1343	38.7	675	3	US-09-954-983-7
37	1232.5	35.5	610	4	US-10-670-695-30
38	654	18.9	438	3	US-09-833-245-137
39	652	18.8	438	3	US-09-833-245-138
40	643	18.5	427	4	US-10-108-260A-4490
41	636	18.3	611	4	US-10-670-695-32
42	414.5	12.0	538	3	US-09-902-535-2
43	414.5	12.0	538	4	US-10-114-893-135
44	414.5	12.0	538	4	US-10-016-249-4
45	413.5	11.9	538	4	US-10-133-036-2

ALIGNMENTS

RESULT 1

US-09-851-859A-6
; Sequence 6, Application US/09851859A
; Patent No. US20020065407A1
; GENERAL INFORMATION:
; APPLICANT: Banerjee, Papia T.
; APPLICANT: Patience, Clive
; APPLICANT: Anderson, Goran K.
; TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of Use
; FILE REFERENCE: 61750-321
; CURRENT APPLICATION NUMBER: US/09/851,859A
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/097,015
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: US 09/376781
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PERV-C
; OTHER INFORMATION: polypeptide sequence taken from GenBank Accession
; OTHER INFORMATION: No. US20020065407A1 AF038600 for comparison.
US-09-851-859A-6

Alignment Scores:
Pred. No.: 1-97e-312 Length: 638
Score: 3420.00 Matches: 637
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 98.6% Indels: 0

DB: 3 Gaps: 0
US-10-723-552-3_COPY_5620_7533 (1-1914) x US-09-851-859A-6 (1-638)
QY 1 ATGCATCCACGTTAAACCGCGCCACCTCCGATTCGGGGTGGAAACCCGAAAGACTG 60
Db 1 MethisProThrLeuAsnArgHisLeuProIleArgGlyGlyLeuProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSergIn 40
QY 121 ACTAATGTATGCGCATAGGACAGCCTGAACTCCCATAAACCTTATCTCTCACCTGG 180
Db 41 ThrIleGlyMetArgIleGlyAspSerLeuIleAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAAGTCCGCGCACAGGTATTAATATCAACAACACTCAAGGGAGGCTCTTTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGGCTCAGATCAGTATTTCCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCAGATATCCTCATGCTCAGCGATTTTATGTTTCCAGGACACCAACAAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCGGGAAATCCCGAGAGTTCCTTTTGTAAACAATGGAACCTGTGTAACCTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGATATTTGGAATGGCAACCTCTCAGCGAGTAGGTAAGTTTCTTATGTCAAC 480
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACAGCTCTGGCAATTTAATTAACCTGACCTGATTTAGAACCGCCCAAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCCTCAGACTAGATTACCTAAAAATAAGTTTCACTGAGAAAGGAAACCAAGAA 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTAAATGGGTAAATGGTATGCTTGGGGAATGGTATATTATGGAGGCTCGGCT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyLysGly 220
QY 661 AAACAACAGGCTCCATTTCACTAATTCGCTCAAAATAAACACAGCTGGAGCTCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProPromet 240
QY 721 GCTATAGGACCAATATACGGTCTTGACGGGTCAAAGACCCCAACCAAGGACCCAGGACCA 780
Db 241 AlalieGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
QY 781 TCCTCTAACAATACCTCTGGATCAGACCCCACTGAGTCTTAACAGACAGCACTAAATGGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerSerSerThrThrLysMetGly 280
QY 841 GCAAAATTTTATAGCTCATCCAGGAGCTTTTCAAGCTCTTAAGCTTCAAGCTCCAGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
QY 901 GCTACCTCTTCTGTGTGGCTATGCTTAGCTTCGGGCCCACTTACTATGAAGGAATGGCT 960
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGlyGlyMetAla 320
QY 961 AGAAGAGGGAATTCATGTGCAAAAGAACATAGAGACCAATGACATGGGATCCCAA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTGGAAGGACCTGTCATAGGAAAGGTTCCTCCCA 1080

Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCACACAAACCTTTGTAAACCACTGAAAGCCCTTTAATCAAACTCTGAGAGCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGGTACTCGTGTATGACAGGTGGTGGCATGTAATACTACTGATTAAACCCCTTGTGTTTCC 1200
Db 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTTAAACCAAACTAAAGATTTTTCGCAATTATGGTCCAAATTTGTTCCCGAGTG 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAGCAATCCTTGATGAAATGACTACAGAAATCATCGACAAAAG 1320
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAGAACCATATCTCTGACACTTGTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTA 1380
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGGAAACAGCTGCCCTGCTCACCGGACCCAGCAGCTAGAAACAGGACTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATGTAAACAGAAAGATCTCCAGCCCTAGAAAAATCTGTCACTAACTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTTAACCTCTTCTGAAGTGTAGTGTAGCCCTTGAAGGAGGATGCTGTTTTATGTGGAT 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGGAGGATTTATGTAGTGTAGCCCTTGAAGGAGGATGCTGTTTTATGTGGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
QY 1621 CATTGAGGGCCATCAGAGACTCCATGAAACAGCTTAGAGAAAGGTTGGAGAGCGTCA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGGAAGAGGAAACTACTCAAGGGTGTTCAGGGATGTTCAACAGGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeu 580
QY 1741 GCTACCCCTACTTTCTGCTTTAACAGGACCCCTTAATAGTCTCTCTCTCTTACTCACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATGTATTATTAAACAGTTAATTCCTTTCCTTATTAGAGAACGATTAAGTCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGGTACTTAGACACACTACCAAGCCCTCTAGCAGGAGCTGGCCGC 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 638
RESULT 2
US-10-441-949-32
; Sequence 32, Application US/10441949
; Publication No. US2004011684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32

; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-32

Alignment Scores:

Pred. No.: 3.29e-311 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 2
Best Local Similarity: 99.5% Mismatches: 1
Query Match: 98.2% Indels: 0
DB: 4 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-32 (1-638)

Qy	1	ATGATCCACGTTAAACCGCGCCACCTCCCGATTGGGTTGGAAAGCCGAAAGACTG	60
Db	1	MethisProthrLeuSerArgArgHisleuProleargGlyLysProLysArgLeu	20
Qy	61	AAATCCCTTAAAGCTTCGCCCTCCATCGCGTGGTTCCTTACTCTGTCAATAACCTCTCAG	120
Db	21	LysileProLeuSerPheAlaSerIleAlaItrPheLeuThrLeuSerIleThrSerGln	40
Qy	121	ACTAATGGTATGCCATAGGAGACGCTGAACTCCCATAAACCTTATCTCTCACTCG	180
Db	41	ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp	60
Qy	181	TTAATTAAGTCTCCGCGCACAGGTATTAATATCAACACACTCAAGGGGGCTCCTTA	240
Db	61	LeuIleThrAspSerGlyThrGlyLeAsnIleAsnAsnThrGlnGlyGluAlaProLeu	80
Qy	241	GGAACTGGTGGCTGTATCTATACGTTTTCCTCAGATCAGTATTCCTAGTCTGACCTCA	300
Db	81	GlyThrTrpTrpProAspLeuTrpValCysLeuArgSerValIleProSerLeuThrSer	100
Qy	301	CCCCCAGATATCTCCATGCTACGGATTTTATATTTGGCCAGGACCAACCAATAATGGA	360
Db	101	ProProAspIleLeuHisAlaHisGlyPheTrpValCysProGlyProProAsnAsnGly	120
Qy	361	AAACATTCGGAATCCAGAGATTTCTTTTGTAAACAATGGAACTGTGTAACTCTAAT	420
Db	121	LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn	140
Qy	421	GATGGATATTGGAAATGGCAACCTCTCAGCAGGATAGGGTAAGTTTCTTATGTCAAC	480
Db	141	AspGlyTrpTrpLysTrpProThrSerGlnAspArgValSerPheSerTrpValAsn	160
Qy	481	ACTATACAGCTCTGGACATTTAATTAATCTGACCTGGATTAAGAACTGGAAGCCCAAG	540
Db	161	ThrTrpThrSerSerGlyGlnPheAsnTrpLeuThrTrpIleArgThrGlySerProLys	180
Qy	541	TGCTCTCTTCAGACCTAGATTACTTAAATAAGTTTCACTGGAAGGAAACACAGAA	600
Db	181	CysSerProSerAspLeuAspTrpLeuLysIleSerPheThrGluLysGlyLysGlnGlu	200
Qy	601	AATATCTTAAATGGGTAAATGGTATGCTTTGGGGAATGGTATATTATGAGGCTCGGGT	660
Db	201	AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTrpTrpGlyGlySerGly	220
Qy	661	AAACAAACGCTCCATTCTAATTCGGCTCAAAATAAACACAGCTGGAGCCTCCCAATG	720
Db	221	LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet	240
Qy	721	GCTATAGGACCAATACGGTCTTGACGGGTCAAGACCCCAACCCCAAGACACAGACCA	780
Db	241	AlaIleGlyProAsnThrValLeuThrGlyGlnArgProThrGlnGlyProGlyPro	260

Qy	781	TCCTCTAACATAACTTCTGGATCAGACCCCACTGAGTCTAACAGCAGCACTAAAAATGGG	840
Db	261	SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly	280
Qy	841	GCAAACTTTTAGCCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCACACTCCAGAG	900
Db	281	AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu	300
Qy	901	GCTACCTCTTCTGCTGCTATGCTTAGCTTCGGGCCCACTTACTACTATGAAGGAATGCT	960
Db	301	AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProTrpTrpGlyGlyMetCala	320
Qy	961	AGAAGAGGAAATTCATATGTGACAAAAGACATAGAGACCAATGCACATGGGGATCCCA	1020
Db	321	ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln	340
Qy	1021	ATAAGCTTACCTTACTGAGGTTCTGGAAAGGACCTGCATAGAGAAAGTTCCCCCA	1080
Db	341	AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro	360
Qy	1081	TCCCAACACACCTTTGTAAACCACTGAAGCCTTTAATCAAACTCTGAGAGTCAATAT	1140
Db	361	SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTrp	380
Qy	1141	CTGGTACCTGGTTATGACAGGTGGTGGCATGTAATACTGGATTAAACCCCTTGTGTTTC	1200
Db	381	LeuValProGlyTrpAspAsgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer	400
Qy	1201	ACCTGGTGGTTTAAACCAACTAAAGATTTTGCATATGGTCCAAATGTTTCCCGAGTG	1260
Db	401	ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal	420
Qy	1261	TATTACTATCCCGAAAGCAATCCTTTGATGATATGACTACAGAAATCATCGACAAAG	1320
Db	421	TyrTrpTrpProGluLysAlaIleLeuAspGluTrpAspTrpArgAsnHisArgGlnLys	440
Qy	1321	AGAAACCCATATCTCTGACACTTGTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTA	1380
Db	441	ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGlyVal	460
Qy	1381	GGAAACAGGAACAGCTGCCCTGGTCAGGGACCAAGCAGCAGTAGAAAACAGACTTAGTA	1440
Db	461	GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn	480
Qy	1441	CTACATCGAATCTTAACAGAGATCTCCAGCCCTAGAAAATCTGCTAGTAACTCGAG	1500
Db	481	LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu	500
Qy	1501	GAATCCCTAACCTCTTATCTGAAGTAGTCTCAGAAATAGAGAGGTTTAGATTATTATA	1560
Db	501	GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu	520
Qy	1561	TTTCTTAAAGAGAGAGATTATGTGTAGCTTGAAGAGGAATGCTGTTTTTATGTGAT	1620
Db	521	PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluCysCysPheTrpValAsp	540
Qy	1621	CATTACGGGCCCATCAGAGCTCCATCAACCAAGCTTAGAGAAAGTTGGAGAGCGTCGA	1680
Db	541	HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGlyLysArgArg	560
Qy	1681	AGGGAAGAGAACTACTCAAGGCTGTTTGGAGGATGGTTCAACAGGCTCTCTTTGGTGT	1740
Db	561	ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu	580
Qy	1741	GCTACCTCTTCTCTTTAACAGGACCTTAAATAGTCTCTCTCTCTCTCTCTCTCTCAG	1800
Db	581	AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal	600
Qy	1801	GGCCCATGTATTATTAAAGTTAATGCTTCAATTAGAGAACGAATAGTGCAGTCCAG	1860
Db	601	GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln	620

```
Qy 1861 ATCATGGTACTTAGACAACAGTACCAAGCCGCTCTAGCAGGAAGCTGGCCGC 1914
Db 621 IlemetValLeuArgGlnGlnTrpGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 3
US-10-441-949-34
; Sequence 34, Application US/10441949
; Publication No: US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodriago, Allen
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-34

Alignment Scores:
Pred. No.: 3 29e-311 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: 4 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-34 (1-638)
Qy 1 ATGCATCCACCGTTAAACCGCGCCACCTCCGATTCGGGTGGAAAGCCGAAAGACTG 60
Db 1 MethisProThrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
Qy 61 AAAATCCCTTAAGTTCGCTCCATCGCGTGTTCCTTACTCTGCTCAATCAACTCTCAG 120
Db 21 LysileProLeuSerPheAlaSerIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 40
Qy 121 ACTAATGATGCGCATAGGAGACAGCCTGAACTCCCATAAACCTTATCTCTCACTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
Qy 181 TTAATTAAGTCTCGGCACACAGGTATTAATATCAACACACTCAAGGGGAGGCTCTTTA 240
Db 61 LeuileThrAspSerGlyThrGlyLeuAsnIleAsnIleAsnIleAsnIleAsnIleAsn 80
Qy 241 GGAACCTGGTGGCTGATCTATAGCTTTGCTTCAGATCAGTATTCCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTrpValCysLeuArgSerValIleProSerLeuThrSer 100
Qy 301 CCCCAGATATCTCCATGCTCACGATTTTATGTTTGGCCAGGACACCAACAAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTrpValCysProGlyProProAsnAsnGly 120
Qy 361 AAACATTGCGGAAATCCAGAGATTTCTTTTGTAAACAATCGAACTGTGTAACTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPheCysLysGlnTrpAsnCysValThrSerAsn 140
Qy 421 GATGATATTCGAATGCGCAACCTCTCAGCAGGATAGGTTAGTTTCTTATGTCAC 480
Db 141 AspGlyTrpTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTrpValAsn 160
Qy 481 ACCTATACCAGCTCTGGACAATTTAATTAACCTGACCTGGATTAGAACTGGAAAGCCCAAG 540
```

```
Db 161 ThrTyThrSerSerGlyGlnPheAsnTyLeuThrTrpIleArgThrGlySerProLys 180
Qy 541 TGCTCTCTCTCAGACCTAGATTACCTAAATAAATAGTTTCTACGTGAGAAAGGAAACAGAA 600
Db 181 CysSerProSerAspLeuAspTyLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
Qy 601 AATATCTTAAATGGGTAAATGGTATGTCCTGGGAATGGTATATATATGGAGGCTCGGCT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyTrpGlyGlySerGly 220
Qy 661 AAACAACACGAGCTCCATTCTAAGCTATTCGGCTCAAAATAAAACACAGCTGGAGCTCCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProMet 240
Qy 721 GCTATAGGACCAATACGCTCTTGACGGGTCAAGACCCCAACCAAGACACGAGACCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProThrGlnGlyProGlyPro 260
Qy 781 TCCTCTAACATACTCTGGATCAGACCCCACTGAGTCTAACACAGCAGCAGCTTAAATGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
Qy 841 GCAAACTTTTGTAGCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCAGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
Qy 901 GCTACCTCTCTCTGCTATGCTTAGCTTCGGGCCCACTTACTATGAAGGATGGCT 960
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProTrpTyTrpGlyGlyMetAla 320
Qy 961 AGAAGAGGGAATTCATATGACAAAGAACATAGACCAATGCAATGCGATGGGATCCCA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
Qy 1021 AATAAGCTTACCTTACTGAGGTTTCTGGAAAGGACCTGCTCATAGGAAAGGTTTCCCCA 1080
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
Qy 1081 TCCCAACAACACCTTTGTAAACCACTGAAAGCTTTTAACTCAACCTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTy 380
Qy 1141 CTGTACTCTGTTATGACAGTGGTGGCATGTAATACTGATTAACCCCTTGGTTGTTTC 1200
Db 381 LeuValProGlyTyAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
Qy 1201 ACCTTGGTTTTTAAACCAAACTTAAAGATTTTTCATTTATGGTCCAAATTTGTTCCCGAGTG 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
Qy 1261 TATTACTATCCGAAAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAG 1320
Db 421 TyTrpTyTrpProGluLysAlaIleLeuAspGluTyAspTyArgAsnHisArgGlnLys 440
Qy 1321 AGAAGACCATATCTCTGACACTTGCCTGATGCTCGACTCTGAGTGGGAGGAGAGTGA 1380
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGlyVal 460
Qy 1381 GGAACAGGAACAGCTGCCCTGGTCCAGGGACCAACAGCAGCTAGAAAACAGGACTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
Qy 1441 CTACATCGAATTTAACAAGAGATCTCAAGCCCTAGAAAAATCTGTCACTTAACCTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
Qy 1501 GAATCCCTTAACCTCTTATCTGAAGTAGTCTCAGAGATAGAGAGGAGGTAGATTTATTA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
Qy 1561 TTTCTAAAGAGAGGAGGATTATGTGTAGCCCTTGAAGGAGGAATCTGTTTTTATGTGGAT 1620
```

```
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysPheTyrValAsp 540
Qy 1621 CATTACGGGGCCATCAGAGACTCCATGAACAAGCTTAGAGAAAGTTGGAGAAGCGTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
Qy 1681 AGGGAAGGAACCTACTCAAGGTGGTTTGAGGATGGTTCAACAGGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
Qy 1741 GCTACCCCTACTTCTGCTTTTAACAGGACCTTAATAGTCTCTCTCTCTACTCACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrVal 600
Qy 1801 GGGCCATGTATTATTAAACAAGTTAATTCCTCTTCAATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
Qy 1861 ATCATGGTACTTAGACAACAGTACCAAGCCCGTCTAGCAGGGAAGCTGGCCGC 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerSerArgGluAlaGlyArg 638

RESULT 4
; Sequence 40, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-40

Alignment Scores:
Pred. No.: 3,29e-311 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: 4 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-40 (1-638)
Qy 1 ATGCATCCACGGTTAAACCGCGCCACCTCCCGATTCGGGTGGAAGCGGAAAGACTG 60
Db 1 MethisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu 20
Qy 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCATAAACCCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
Qy 121 ACTAATGGTATGGCATAGGAGACAGCTGAACTCCCAATAAACCCCTTATCTCAACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTip 60
Qy 181 TTAATTAAGTACCTCCGACAGGTATTAAATATCAACAAGTCTAGGGGAGGCTCCCTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyIleAsnIleAsnThrGlnGlyGluAlaProLeu 80
```

```
Qy 241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTATTCTTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
Qy 301 CCCCAGATATCTCTCCATGCTCAGGATTTTATGTTGCCCGCAGGACCAACCAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
Qy 361 AAACATTCGCGAAATCCACAGAGATTTCTTTTGTAAACAATGGAACCTGTGTAACCTCAAT 420
Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
Qy 421 GATGGATATTGGAAATGGCAACCTCTCAGCAGATAGGGTAAGTTTCTTATCTCAAC 480
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn 160
Qy 481 ACCTATACAGCTCTGGACAAATTAATTAATCTGACCTGGATTAGAACTGGAAGCCCCAAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
Qy 541 TGCTCTCTCTCAGACCTAGATTACCTAAAATAAGTTTCACTGAGAAGGAAACAGAA 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
Qy 601 AATATCCTAAATGGGTAAATGGTATGCTCTTGGGGAATGGTATATTATGAGAGGCTCGGGT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGly 220
Qy 661 AAACAACAGGCTCCATTTCAATTAATTCGCTCAAAATAAACACAGCTGGAGCTCCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
Qy 721 GCTATAGGACCAATACGGTCTTGACGGGTCAAGACCCCAAGCCCAAGGACGAGGACCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProThrGlnGlyProGlyPro 260
Qy 781 TCCTCTAACATAACTCTCTGGATCAGACCCCACTGAGTCTAACACAGCAGCTATAAATGGG 840
Db 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
Qy 841 GCAAAACCTTTTAGCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAG 900
Db 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
Qy 901 GCTACCTCTTCTTGGCTATGCTTAGCTTCGGGCCACCTTACTACTATGAGGAATGGCT 960
Db 301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGlyGlyMetAla 320
Qy 961 AGAAGAGGGAATTTCAATGTGACAAAAGAACATAGAGACCAATGCACATGGGATCCCAA 1020
Db 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
Qy 1021 AATAAGCTTACCTTACTAGAGTTTCTGGAAGGACCTGCATAGAAAGGTTCCCCCA 1080
Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
Qy 1081 TCCCACCAACACCTTTGTAAACACACATGAAGCCTTTAATCAACCTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
Qy 1141 CTGGTACCTGGTTATGACAGGTGGTGGCATGTAACTAGTAACTGATTAACCCCTTGTTTCC 1200
Db 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
Qy 1201 ACCTTGGTTTTTAAACCAAACTAAAGATTTTTCATTATGGTCCAAATTTGTTCCCGAGT 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
Qy 1261 TATTACTATCCCGAAAAAGCAATCCTTGATGATATGACTACAGAAATCATCGCAAAAG 1320
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
```

QY 1321 AGAGAACCCATATCTCTGACACTTGCTGCTGCTCGGACTTGGAGTGGCAGCAGGTGTA 1380
Db |||||
QY 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
Db |||||
QY 1381 GGAAACAGGAAACAGCTCGCTGCTACGCGGACACACAGCAGCTAGAAACAGGACTTAGTAAC 1440
Db |||||
QY 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
Db |||||
QY 1441 CTACATCGAATGTAAACAGAGATCTCAAGCCCTAGAAAATCTGTCAAGTAACCTGGAG 1500
Db |||||
QY 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
Db |||||
QY 1501 GAACTCCCTAACTCTGCTGAACTAGTCTACAGATAGAGAGGGTGTAGATTATTA 1560
Db |||||
QY 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
Db |||||
QY 1561 TTTCTAAAGAGGAGGATATGTGTAGCCTTGAAGAGGAGTAATGCTGTTTTTGTGGAT 1620
Db |||||
QY 521 PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysCysPheTyrValAsp 540
Db |||||
QY 1621 CATTGAGGGGCATCAGAGATCCATGAACAGCTTAGAGAAAGTTGGAGAGCGTGA 1680
Db |||||
QY 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
Db |||||
QY 1681 AGGGAARAGGAACTACTCAAGGNGTTCAGGGATGTTCAACAGGTCTCTTTGGTTG 1740
Db |||||
QY 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
Db |||||
QY 1741 GCTACCTACTTCTGCTTTAAACAGGACCTTAATAGTCCCTCTCTGTTACTCACAGTT 1800
Db |||||
QY 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
Db |||||
QY 1801 GGGCCATGTATTATTAACAAGTAAATGCTTCAATAGAGAACGAATAGTGCAGTCCAG 1860
Db |||||
QY 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
Db |||||
QY 1861 ATCATGTACTTACACACAGTACCAGACCGCTCTACAGGAGCTGGCGC 1914
Db |||||
QY 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 638

RESULT 5

; Sequence 42, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-42

Alignment Scores:
Pred. No.: 3,298-311 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: 4 Gaps: 0

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-42 (1-638)
QY 1 ATGCATCCCGCTTAAACCGCGCCACCTCCCGATTCCGGGTGGAAGCCGAAAGACTG 60
Db |||||
QY 1 MethIsProThrLeuSerArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
Db |||||
QY 61 AAAATCCCTTAAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTCAATAACCTCTCAG 120
Db |||||
QY 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
Db |||||
QY 121 ACTAATGTATGCTCCATAGGAGACAGCCTGAACCTCCATAAACCTTATCTCTCACCTGG 180
Db |||||
QY 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
Db |||||
QY 181 TTAATTACTGCTCCGCGCACAGGTATTAAATATCAACACTCAAGCGGAGGCTCCTTTA 240
Db |||||
QY 61 LeuIleThrAspSerGlyThrGlyLeuAsnAsnThrGlnGlyGluAlaProLeu 80
Db |||||
QY 241 GGAACCTGGTGGCTGATCTATAGCTTTTGCCTCAGATCAGTTATTTCCTAGTCTGACCTCA 300
Db |||||
QY 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
Db |||||
QY 301 CCCCACAGATATCTCCATGCTCAGCGATTTTATGTTGCCAGGACACCAACAAATAATGGA 360
Db |||||
QY 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
Db |||||
QY 361 AAACATTGCGAAATCCAGAGATTCTTTTGTAAACAATGGAACGTGTAACTCTTAAT 420
Db |||||
QY 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
Db |||||
QY 421 GATGATATTGGAATGGCAACCTCTCAGCAGATAGGTAAAGTTTCTTATGTCTAAC 480
Db |||||
QY 141 AspGlyTyrTrpLysTrpProThrSerGlnAspArgValSerPheSerTyrValAsn 160
Db |||||
QY 481 ACCTATACCACTCTGGACAAATTTAATACCTGCTGAGTGTAGAACTGGGAAGCCCAAG 540
Db |||||
QY 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
Db |||||
QY 541 TGCTCTCTTCAGACCTAGATTACTTAAATAAGTTTCACTGAGAAAGGAAACAAAGAA 600
Db |||||
QY 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
Db |||||
QY 601 AATATCTTAAATGGTAAATGTTATCTCTTGGGGAATGGTATATTATGGAGGCTCGGT 660
Db |||||
QY 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlySerGly 220
Db |||||
QY 661 AAACAACCAAGCTCCATTCTAACTATTCCCTCAAAAATAAACCCAGCTCGAGCTCCAATG 720
Db |||||
QY 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
Db |||||
QY 721 GCTATAGACCAATATACGTTCTGACGGGTCAAGAGACCCCAACCCCAAGGACCAGACCA 780
Db |||||
QY 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyPro 260
Db |||||
QY 781 TCCTCTAACATAACTTCTGGATCAGACCCACTGAGTCTACAGCACCACTAAATATGGG 840
Db |||||
QY 261 SerSerAsnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
Db |||||
QY 841 GCAAAACCTTTTATAGCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCAGCACTCCAGAG 900
Db |||||
QY 281 AlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGlu 300
Db |||||
QY 901 GCTACCTCTTCTGTGTGGCTAGTTAGCTTCGGGCCCACTTACTTACTATGAGGAATGGCT 960
Db |||||
QY 301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAla 320
Db |||||
QY 961 AGAAGAGGGAATTCATGTGACAAAGAAACATAGAGACCAATGCACATGGGATCCCAA 1020
Db |||||
QY 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
Db |||||
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTGGAAAAAGCACCTCATAGGAAGAGGTTCCTCCA 1080
Db |||||

Db 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
QY 1081 TCCCAACCAACCTTTGTAACCACTGAAGCTTTAAATCAAACTCTGAGAGTCAATAT 1140
Db 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
QY 1141 CTGGTACCTGGTTATGACAGAGTGGTGGCATGTAACTACTGATTAACCCCTTGTTTCC 1200
Db 381 LeuValProGlyTyrAspArgTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
QY 1201 ACCTTGGTTTAAACCAAACTAAAGATTTTGCTATTATGGTCCAAATTTGTTCCCGAGTG 1260
Db 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
QY 1261 TATTACTATCCGAAAAAGCAATCTTCGATGAATATGACTACAGAAATCATCGACAAAG 1320
Db 421 TyrTyrTyrProGluLysAlaIleLeuAspGlnTyrAspTyrArgAsnHisArgGlnLys 440
QY 1321 AGAGAACCATATCTTCACACTTCTGCTGTGATGCTCGCACTTGGAGTGGCAGCGTGA 1380
Db 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
QY 1381 GGAACAGGAACAGCTGCCCTGTGTACGGGACACACAGCAGCTAGAAACAGAGCTTAGTAAC 1440
Db 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
QY 1441 CTACATCGAATTTGTAACAGAGATCTCCAAGCCCTAGAAAAATCTGTCAAGTAACCTGGAG 1500
Db 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
QY 1501 GAATCCCTAACCTCTTATCTGAGTAGTCTCTACAGAAATGAGAGGGTTAGATTATTA 1560
Db 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
QY 1561 TTTCTAAAGAGAGGATTTATGTAGCTTGTAGGAGGAATGCTGTTTTTATGTGAT 1620
Db 521 PheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGlyCysPheTyrValAsp 540
QY 1621 CATTGAGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAAGGTTGAGAAAGCTCGA 1680
Db 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
QY 1681 AGGAAAGAAACTACTCAAGGGTGGTTTGAAGGATGGTTCAACAGGTCTCTTTGGTTG 1740
Db 561 ArgGluLysGluThrThrGlnGlyTyrPheGluGlyTyrPheAsnArgSerProTyrLeu 580
QY 1741 GCTACCTTACTTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTGTTACTCACAGTT 1800
Db 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
QY 1801 GGGCCATGTATTATTAAACAGTTAAATGGCTTTCATTAGAGAACGAATAAGTGCAGTCCAG 1860
Db 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
QY 1861 ATCATGTACTTAGACACAGTACCAAGCCCGTCTAGCAGGGAAGCTGGCCG 1914
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 638

RESULT 6

US-10-441-949-45
; Sequence 45, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodriago, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659

; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-45

Alignment Scores:
Pred. No.: 3,29e-311 Length: 638
Score: 3407.00 Matches: 635
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 98.2% Indels: 0
DB: 4 Gaps: 0

US-10-723-552-3_copy_5620_7533 (1-1914) x US-10-441-949-45 (1-638)

QY 1 ATGCATCCACGTTTAAACCGGCGCACCTCCCGATTGGGGTGGAAAGCCGAAAGACTG 60
Db 1 MethisProthrLeuSerArgHisLeuProIleargGlyLysProLysargLeu 20
QY 61 AAAATCCCTTAACTTCGCTCCATCGCGTTCCTTACTCTGTCAATAACCTCTCAG 120
Db 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGGTATGCCATAGGAGACAGCTGAACTCCCATAAACCTTATCTCTCACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTCAGTCCCGCACAGGTATTAAATCAACAACACTCAAGGGAGGCTCCTTTA 240
Db 61 LeuIlethrAspSerGlyThrGlyIleAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGTATCTATACGTTTGGCTCAGATCAGTTATTCCTAGTCTGACCTCA 300
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuThrSer 100
QY 301 CCCCCAGATATCTCCATGTCTACGGATTTTATGTTTCCCGAGCACCAAAATAATGGA 360
Db 101 ProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProProAsnAsnGly 120
QY 361 AAACATTCCGGAATCCACAGAGATTTCTTTTGTAAACAATGGAACCTGTGTAACCTCTAAT 420
Db 121 LysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysValThrSerAsn 140
QY 421 GATGGATATTGGAAATGCCAACCTCTCAGCAGGATAGGGTAAAGTTTTTCTTATGTCAC 480
Db 141 AspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSerTyrValAsn 160
QY 481 ACCTATACAGCTCTGGACAAATTAATTAATCCTGACCTGGATTAGAACTGGAAGCCCCAAG 540
Db 161 ThrTyrThrSerSerGlyGlnPheAsnTyrLeuThrTrpIleArgThrGlySerProLys 180
QY 541 TGCTCTCCTTCAGACCTAGATTACCTAAAATAAGTTTCTCACTGAGAAAGGAACAAAGAA 600
Db 181 CysSerProSerAspLeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlu 200
QY 601 AATATCTCAAAATGGGTAAATGGTATGCTTTGGGAATGGTATATATGAGGGCTCGGGT 660
Db 201 AsnIleLeuLysTrpValAsnGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGly 220
QY 661 AAACAACAGGCTCCATTCTAACTATTTCGCTCAAAATAAACACAGCTGGAGCTCCCAATG 720
Db 221 LysGlnProGlySerIleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMet 240
QY 721 GCTATAGGACCAATAGGCTCTTGACGGGTCAAGACCCCCCAAGGAGGACGAGGACCA 780
Db 241 AlaIleGlyProAsnThrValLeuThrGlyGlnArgProThrGlnGlyProGlyPro 260

```
QY 781 TCCTCTAATAACTTCTGGATCAGACCCCACTAGTCTAAACAGCACGACTAAATGGG 840
Db |||||||
QY 261 SerSerAlnIleThrSerGlySerAspProThrGluSerAsnSerThrThrLysMetGly 280
Db |||||||
QY 841 GCAAAACTTTTATAGCCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCACGACTCCAGAG 900
Db |||||||
QY 281 AlalysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAenSerThrThrProGlu 300
Db |||||||
QY 901 GCTACTCTTCTTGTGTGCTATGCTTACGTCGGGCCCACTTACTACTAAGAAGGATGGCT 960
Db |||||||
QY 301 AlaThrSerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAla 320
Db |||||||
QY 961 AGAAGAGGAATTCAAATGTCAGAAACAAACATAGAGACCAATGCACATGGGGATCCCAA 1020
Db |||||||
QY 321 ArgArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGln 340
Db |||||||
QY 1021 AATAAGCTTACCTTACTGAGGTTTCTCGAAAAGCCACTCGATAGGAAAGTTCCCCCA 1080
Db |||||||
QY 341 AsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProPro 360
Db |||||||
QY 1081 TCCACCAACACCTTTGTAAACACACTGAAGCTTTTAATCAAACTCTGAGAGTCAATAT 1140
Db |||||||
QY 361 SerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyr 380
Db |||||||
QY 1141 CTGGTACCTGTTATGACAGGTGGTGGCATGTAATCTGATTAAACCCCTTGTTGTTCC 1200
Db |||||||
QY 381 LeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSer 400
Db |||||||
QY 1201 ACCTTGGTGTTTTAAACAACTAAAGATTTTTCGATTATGCTCCAAATTTGTTCCCGAGTG 1260
Db |||||||
QY 401 ThrLeuValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgVal 420
Db |||||||
QY 1261 TATTACTATCCGAAAAGCAATCTTGATGATATGACTACAGAAATCATCGACAAAG 1320
Db |||||||
QY 421 TyrTyrTyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLys 440
Db |||||||
QY 1321 AGAGAACCATATCTCTCACACTGCTGATGCTCGACTTGGAGTGGCAGCAGGTGTA 1380
Db |||||||
QY 441 ArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyVal 460
Db |||||||
QY 1381 GAAACAGGAAACAGCTCGCTGTGTCAGGGACACACAGCAGCTAGAAACAGGACTTAGTAAC 1440
Db |||||||
QY 461 GlyThrGlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsn 480
Db |||||||
QY 1441 CTACATCGAANTGTAAACAGAGATCTCCAGCCCTAGAAAAATCTGTGACGTAACCTGAG 1500
Db |||||||
QY 481 LeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGlu 500
Db |||||||
QY 1501 GAATCCCTTAACCTCTTATCTGAAGTACTCTACAGAAATAGAGAGGTTAGATTATTA 1560
Db |||||||
QY 501 GluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeu 520
Db |||||||
QY 1561 TTTCTAAAGAGAGGAGGATATGTCTAGCCTTGAAGAGGAATGCTGTTTTTATGTGGAT 1620
Db |||||||
QY 521 PheLeuLysGluGlyLeuCysValAlaLeuLysGluGluCysPheTyrValAsp 540
Db |||||||
QY 1621 CATTCAGGGGCATCAGAGACTCCATGAACAGCTTAGAGAAAGTTGGAGAGGCTCGA 1680
Db |||||||
QY 541 HisSerGlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArg 560
Db |||||||
QY 1681 AGGGAAGGAACTACTCAAGGNGTGTGAGGATGTTCAACAGGTCTCTTTGGTTG 1740
Db |||||||
QY 561 ArgGluLysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeu 580
Db |||||||
QY 1741 GCTACCTTACTTTTCAAGGACCTTAAATAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 1800
Db |||||||
QY 581 AlaThrLeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuThrVal 600
Db |||||||
QY 1801 GGGCCATGATATTATTAACAAGTTAATTCCTTCATTAGAGAACGAATAGTCAGTCCAG 1860
Db |||||||
QY 601 GlyProCysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGln 620
Db |||||||
QY 1861 ATCATGGTACTTTAGACAACAGTACCAAGCCCGTCTAGCAGGGAAGCTGGCGC 1914
Db |||||||
```

```
Db 621 IleMetValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 638
Db |||||||
RESULT 7
US-10-441-949-36
; Sequence 36, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-36

Alignment Scores:
Pred. No.: 8,02e-303 Length: 678
Score: 3318,00 Matches: 634
Percent Similarity: 93.8% Conservative: 1
Best Local Similarity: 93.8% Mismatches: 3
Query Match: 95.7% Indels: 38
DB: 4

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-36 (1-678)
QY 1 ATGCATCCAGCTTAAACCGCGCCACCTCCGATTCCGGTGGAAAGCCCAAGACACTG 60
Db 1 MethisProThrLeuSerArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAAGCTCGCTCCATCGCGGTTCTTACTCTGTCTAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40
QY 121 ACTAATGGTATGCGCATAGGACAGCGCTGAACCTCCCAATAAACCTTATCTCTCACCTGG 180
Db 41 ThrAsnGlyMetArgIleGlyAspSerLeuAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAAGTATGCTCGGCACAGGATTAATATCAACAACACTCAAGGGAGGCTCTCTTA 240
Db 61 LeuIleThrAspSerGlyThrGlyLeuAsnIleAsnAsnThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTTGGCTGCTGATCTATAGCTTTCCTCCATCAGATCAGTATTCTCTAGTCTG 294
Db 81 GlyThrTrpTrpProAspLeuTyrValCysLeuArgSerValIleProSerLeuAsnAsp 100
QY 295 ---ACCTCACCCCGAGATATCTCATCTCACGGATTTATGTTTGGCCAGGACCACCA 351
Db 101 GlnThrSerProProAspIleLeuHisAlaHisGlyPheTyrValCysProGlyProPro 120
QY 352 AATAATGAAACATTTGCGGAAATCCCGAGATTTCTTTTGTAAACAATGAAGTGTGTA 411
Db 121 AsnAsnGlyLysHisCysGlyAsnProArgAspPhePheCysLysGlnTrpAsnCysVal 140
QY 412 ACCTCTAATGATGATATTGGAATGGCCAACTCTCAGCAGGATAGGTTAGTTTCT 471
Db 141 ThrSerAsnAspGlyTyrTrpLysTrpProThrSerGlnGlnAspArgValSerPheSer 160
QY 472 TATGTCAACACTATACCAGCTCTGGACAATTTAATTAC-----CTGACC 516
Db |||||||
```


Db 161 TyrValAsnThrThrThrSerSerGlyGlnPheAsnThrGlyHisGlyArgTrpLeuThr 180
Qy 517 TGG-----ATTAGAACTGGAGCCCAAGTGTCTCTCTTCAGAC 555
Db 181 TrpGlnGlnArgValGlnLysAspIleArgThrGlySerProLysCysSerProSerAsp 200
Qy 556 CTAGATTACCTAAATAAGTTTCACTCAGAAAGGAAACAAGAAATATCTTAAATGG 615
Db 201 LeuAspThrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleLeuLysTrp 220
Qy 616 GTAATAGTATGCTTGGGGGAATGGTATATTATGAGGCTCGGTAAACAACACCGGCTCC 675
Db 221 ValAsnGlyMetSerTrpGlyMetValTyrGlyGlySerGlyLysGlnProGlySer 240
Qy 676 ATTCTAATATTCGCTCAAAATAAAC---CAGCTGGAGCCTCCAATGGCTATAGGACCA 732
Db 241 IleLeuThrIleArgLeuLysIleAsnThrGlnLeuGluProProMetAlaIleGlyPro 260
Qy 733 AATAGGCTCTGAGGGTCAAGAGCCCAACCCCAAGGACCA-----774
Db 261 AsnThrValLeuThrGlyGlnArgProProThrGlnGlyProProHisAsnLeuProVal 280
Qy 775 -----GGACCATCTCT-----AACATAACTTCTGGA 801
Db 281 ProGlnGlyProSerProAsnProAspIleThrGlnSerAspTyrAsnIleThrSerGly 300
Qy 802 TCAGACCCCACT-----GAGTCTAACAGCAGCACTAAATGGGGCAAAA 846
Db 301 SerAspProThrAsnThrProArgAsnGluSerAsnSerThrThrLysMetGlyAlaLys 320
Qy 847 CTTTTAGCTCATCGAGGAGCTTTCAAGCTCTTAACCTCCAGCACTCCAGAGGCTACC 906
Db 321 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThr 340
Qy 907 TCTCTTTGGCTATGCTTACCTTCGGGCCCACTTACTATGAAGAAATGGCTAGAGA 966
Db 341 SerSerCysTrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAlaArg 360
Qy 967 GGGAAATTCATGTGACAAAAGACATAGAGACCAATGCATGGGATCCCAAAATAAG 1026
Db 361 GlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLys 380
Qy 1027 CTTACCTTACTGAGGTTCTGGAAGGCACTGCATAGAAAGTTCCCCCACTCCAC 1086
Db 381 LeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHis 400
Qy 1087 CAACACCTTGTACACCACTGAGCCTTTAATCAACCTCTGAGACTCAATATCTGTA 1146
Db 401 GlnHisLeuCysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyrLeuVal 420
Qy 1147 CCTGGTTATCAGAGTGGTGGCATGTAATCTGATTAACCTCTGTGTTTCCACCTTG 1206
Db 421 ProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeu 440
Qy 1207 GTTTTTAACCAACTAAAGATTTTGCATTATGTCCTCAAAATGTTCCCGAGTGATPAC 1266
Db 441 ValPheAsnGlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrTyr 460
Qy 1267 TATCCCGAAGCAATCCTTGATGATATGATACAGCAATCATCGCAAAAGAGGAA 1326
Db 461 TyrProGluLysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLysArgGlu 480
Qy 1327 CCATATCTCTGACACTTGTCTGATCTCGGACTTGGAGTGGCAGCAGGTGAGGAACA 1386
Db 481 ProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThr 500
Qy 1387 GGAACAGCTGCCCTGGTGCAGGACCAACAGCAGCTAGAAAACAGACTTAGTACCTACAT 1446
Db 501 GlyThrAlaAlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHis 520
Qy 1447 CGAATTGTACAGAGATCTCAAGCCCTAGAAAAATCTGTACAGTACCTGGAGGATCC 1506
Db 521 ArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSer 540

Qy 1507 CTAACCTCCTTATCTGAAGTAGTCTCTACAGATAGAGGGTTAGATTATTATTCTTA 1566
Db 541 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 560
Qy 1567 AAAGAACGAGGATTTATGTAGCTTGAAGAGGAATGCTGTTTTATGTGGATCATTCATCA 1626
Db 561 LysGluGlyGlyLeuCysValAlaLeuLysGluCysCysPheTyrValAspHisSer 580
Qy 1627 GGGCCCATCAGAGACTCATGAACAAGCTTAGAGAAAGTTGGAGAGCGTCGAAGGAA 1686
Db 581 GlyAlaIleArgAspSerMetAsnLysLeuArgGluArgLeuGluLysArgArgGlu 600
Qy 1687 AAGGAAACTACTCAAGGGTGTGTGAGGATGGTTCAACAGGCTCTTTTGGTTGGCTACC 1746
Db 601 LysGluThrThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeuAlaThr 620
Qy 1747 CTACTTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTGTTACTCAGAGTGGGCA 1806
Db 621 LeuLeuSerAlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuValGlyPro 640
Qy 1807 TGTATTATTAAACAAGTTAATTGCTTCATTAGAGAACGAATAAGTCAGTCCAGATCATG 1866
Db 641 CysIleIleAsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMet 660
Qy 1867 GTACTTAGACAACAGTACCAAGCCCTCTAGCAGGGAAGCTGGCCGC 1914
Db 661 ValLeuArgGlnGlnTyrGlnSerProSerArgGluAlaGlyArg 676

RESULT 8

US-10-441-949-38
; Sequence 38, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-38

Alignment Scores:
Pred. No.: 8, 02e-303 Length: 678
Score: 3318.00 Matches: 634
Percent Similarity: 93.9% Conservative: 1
Best Local Similarity: 93.8% Mismatches: 3
Query Match: 95.7% Indels: 38
DB: 4 Gaps: 7

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-38 (1-678)

Qy 1 ATGCATCCCAAGCTTAACCGCGCCACCTCCCATTCGGGGTGGAAAGCGAAAGACTG 60
Db 1 MetHisProThrLeuSerArgArgHisLeuProIleArgGlyLysProLysArgLeu 20
Qy 61 AAAATCCCTTAGCTTCGGCTCCATCGGTGTTCTTACTCTGTCAATAAAGCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrSerGln 40

; LENGTH: 653
; TYPE: PRT
; ORGANISM: Viral
US-10-029-656-2

Alignment Scores:

Score: 8,83e-287 Length: 653
3147.50 Matches: 593
Percent Similarity: 93.6% Conservative: 18
Best Local Similarity: 90.8% Mismatches: 27
Query Match: 90.8% Indels: 15
DB: 4 Gaps: 3

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-029-656-2 (1-653)

```
QY 1 ATGCATCCACCTTAAACCGGCGCCACCTCCCGATTCCGGGGTGGAAAGCCGAAAGACTG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MethHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyProIleArgLeu 20

QY 61 AAATCCCTTAAGCTTCGCCCTCCATCGCGTTCCTTACTCTCTCAATAACCTCTCAG 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 LysIleProLeuSerPheAlaSerIleAlaIlePheLeuThrLeuSerIleThrProGln 40

QY 121 ACTAATGTATGCGCATAGGACAGACCTGAACTCCCATAAACCTTATCTCTACCTGG 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41 ValAsnGlyLysArgLeuValAsnSerProAsnSerHisLysProLeuSerLeuThrTrp 60

QY 181 TTAATTACTGACTCCGGCACAGGTATTAAATATCAACAACACTCAAGGGGAGCTCTTTA 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80

QY 241 GGAACTGGTGGCTGTATATACCTTTTGGCTCAGATCAGTTATTCCTAGTCTG----- 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100

QY 295 ---ACCTCACCCCGAGATATCTCCATGCTCAGGATTTTATGTTTGGCCAGGACCA 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 GlnAlaThrProProAspValLeuArgAlaIleGlyPheTyrValCysProGlyProPro 120

QY 352 AATAATGAAACATTTGCGGAAATCCAGAGATTTCTTTTGTAAACATGAACGTGTGA 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140

QY 412 ACCTTAATGATGATATGGAATGGCCAACTCTCAGCAGGATAGGTTAAGTTTCT 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnAspArgValSerIleSer 160

QY 472 TATGTCAACACTATACCACTCTGGACAATTTAATAC-----CTGACC 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180

QY 517 TGG-----ATTAGAACTGGAAAGCCCAAGTCTCTCTCTCAGAC 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TrpGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200

QY 556 CTAGATTACCTAAAAATAAGTTTCACTGAGAAAGAAACCAAGAAATATCTTAAATGG 615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleLeuLysTrp 220

QY 616 GTAAATGTATGCTTTGGGGAATGGTATATTATGAGGCTCGGTTAAACACAGGCTCC 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 ValHisGlyMetSerTrpGlyMetValTyrTyrGlyGlySerGlyLysGlnProGlySer 240

QY 676 ATTCTTAATATTCGCTCAAAATAAACCGCTGAGCTCCCAATGGCTATAGGACCAAT 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 IleLeuThrIleArgLeuLysIleAsnGlnLeuGluProProMetAlaIleGlyProAsn 260

QY 736 ACGGTCTTGAGGGTCAAAGACCCCAACCAAGACGAGACCATCTCTTAACATAACT 795
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 ThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyThrSerSerAsnIleThr 280

QY 796 TCTGGATCAGACCCCACTGAGTCTAACAGACGACTAAATGGGGGCAAACTTTTATAGC 855
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 281 SerGlySerAspProThrGluSerAsnSerThrThrLysMetGlyAlaLysLeuPheSer 300
QY 856 CTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCAGACTCCAGAGCTACCTCTTCTTGT 915
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 LeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThrSerSerCys 320

QY 916 TGGCTATGCTTAGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAGAGGGAATTC 975
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 TrpLeuCysLeuAlaLeuGlyProProTyrTyrGluGlyMetAlaArgArgGlyLysPhe 340

QY 976 AATGTGACAAAAGAACAATAGACCAATGACATCGGATGCCAAAATAAGCTTACCTTT 1035
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 AsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLysLeuThrLeu 360

QY 1036 ACTGAGGTTCTCGAAAAAGGCACCTGCATAGAAAGTTCCCCATCCACCAACACCTT 1095
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 ThrGluValSerGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu 380

QY 1096 TGTAAACCACTGAAGCCCTTTAATCAAACTCTGAGACTCAATATCTGTACTCTGTTAT 1155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 CysAsnHisThrGluAlaPheAsnGlnThrSerGluSerGlnTyrLeuValProGlyTyr 400

QY 1156 GACAGTGGTGGCGATGTAATACCTGGAATAACCCCTGTGTTCACCTTGGTGTTTTAA 1215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn 420

QY 1216 CAAACTAAAGATTTCGATTATGTCCTCAAAATGTCCTCCAGTGTATTACTATCCCGAA 1275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 GlnThrLysAspPheCysIleMetValGlnIleValProArgValTyrTyrProGlu 440

QY 1276 AAAGCAATCTCTGATGATATGACTACAGAAATCATCGACAAAAGAGAGAACCCATATCT 1335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 LysAlaIleLeuAspGluTyrAspTyrArgAsnHisArgGlnLysArgGluProIleSer 460

QY 1336 CTGACACTGTCTGTGATGCTCGGACTTGGAGTGGCAGCAGGTAGGAAACAGGAACAGCT 1395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThrGlyThrAla 480

QY 1396 GCCCTGGTCACGGACCAACAGCAGCAGTAAACAGAGCTTAGTACCTACCTCGAATCTGA 1455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 AlaLeuValThrGlyProGlnGlnLeuGlnThrGlyLeuSerAsnLeuHisArgIleVal 500

QY 1456 ACAGAAGATCTCCAAGCCCTAGAAAAATCTCTCAGTAACCTGGAGGAATCCCTAACCTCC 1515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluSerLeuThrSer 520

QY 1516 TTATCTGAAGTAGTCTTACAGATAGAGAGGGTTAGATTATTATTCTTAAAGAAGGA 1575
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 521 LeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuLysGluGly 540

QY 1576 GGATTATGTAGCTTGAAGGAGGAATGCTGTTTTTATGTTGATCATTCAGGGGCCATC 1635
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 GlyLeuCysValAlaLeuLysGluGlyCysPheTyrValAspHisSerGlyAlaIle 560

QY 1636 AGACACTCCATGAACAGCTTAGAGAAAGGTTGGAGAAAGCGCTCAAGGAAAAAGAAACT 1695
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 ArgAspSerMetSerLysLeuArgGluArgLeuGluLysArgArgArgGluLysGluThr 580

QY 1696 ACTCAAGGTGGTTGAGGATGGTTCAACAGGTCTCTTTGGTTGGCTACCTACTTTCT 1755
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 581 ThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerLeuTrpLeuAlaThrLeuLeuSer 600

QY 1756 GCTTTACAGGACCTTAAATAGTCTCTCTCTGTACTACAGTTGGGCGCATGTATTATT 1815
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 AlaLeuThrGlyProLeuIleValLeuLeuLeuLeuLeuThrValGlyProCysIleIle 620

QY 1816 AACAAAGTAAATGGCTCTTATTAGAGAACAATAGTCAGTCCAGATCATCGTACTTAGA 1875
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 AsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMetValLeuArg 640

QY 1876 CAACAGTACCAAGCCCGTCTTAGCAGGAAAGCTGGCCGC 1914
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 641 GlnGlnTyrGlnSerProSerArgSerArgGluAlaGlyArg 653
```

```
RESULT 10
US-10-029-656-4
; Sequence 4, Application US/10029656
; Publication No. US20030175904A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; APPLICANT: Oldmixon, Beth
; APPLICANT: Ericsson, Thomas
; TITLE OF INVENTION: Molecular Sequence of Pig Endogenous Retrovirus Receptor and Meth
; FILE OF INVENTION: Use
; FILE REFERENCE: 329579-3
; CURRENT APPLICATION NUMBER: US/10/029,656
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/60/285,103
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Viral
US-10-029-656-4

Alignment Scores:
Pred. No.: 1,196-285 Length: 653
Score: 3135.50 Matches: 591
Percent Similarity: 93.1% Conservative: 17
Best Local Similarity: 90.5% Mismatches: 30
Query Match: 90.4% Indels: 15
DB: 4 Gaps: 3

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-029-656-4 (1-653)
QY 1 ATGCATCCACGTTAAACGGCGCCACCTCCGATTTCGGGGTGGAAAGCCGAAAGACTG 60
DB 1 MethHisProThrLeuSerArgArgHisLeuProIleArgGlyGlyLeuProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCGTGGTTCCTTACTCTGTCAATACCTCTCAG 120
DB 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATGTATGCGCATAGGACAGCTGAACTCCCATAAACCTTATCTCTCACCTGG 180
DB 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAAGTCTCGGCACAGGTATTAAATATCAACACACTCAAGGGAGGCTCCTTTA 240
DB 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTTATTCTTAGTCTG----- 294
DB 81 GlyThrTrpTrpProGluLeuTrpValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 295 ---ACCTCACCCAGATATCTCCATCTCAGGATTTTATGTTTGGCCAGGACCACCA 351
DB 101 GlnAlaThrProProAspValArgAlaTrpGlyPheTrpValCysProGlyProPro 120
QY 352 AATAATGAAACATTCGCGAAATCCAGAGATTTCTTTTGTAAACAATGGAATGTGTA 411
DB 121 AsnAsnGluGluTrpCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 412 ACCTCTAATGATGGATATGGAAATGGCCACCTCTCAGCAGGATAGGGTAAGTTTCT 471
DB 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTrpSer 160
QY 472 TATGTCAACACCTATACCAAGCTCTCGACAATTTAATTAC-----CTGACC 516
DB 161 PheValAsnAsnProThrSerTrpAsnGlnPheAsnTrpGlyHisGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGAACTGGAAGCCCCCAAGTGCTCTCTTCAGAC 555
DB 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
```

```
QY 556 CTAGATTACCTAAAAATAAGTTTCACTGAGAAAGCAAAACAAGAAAAATATCTTAAATGG 615
DB 201 LeuAspTrpLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleLeuLysTrp 220
QY 616 GTAAATGGTATGTCTTGGGAATGGTATATTATGGAGGCTCGGTAAACAACACAGGCTCC 675
DB 221 ValAsnGlyMetSerTrpGlyMetValTrpGlyGlySerGlyLysGlnProGlySer 240
QY 676 ATTCTAATCTTCCCTCAAAATAAACACAGCTGAGGCTCCAATGGCTATAGGACCAAT 735
DB 241 IleLeuThrIleArgLeuLysIleAsnGlnLeuGluProMetAlaIleGlyProAsn 260
QY 736 ACGGTCTTGCAGGTCAAAAGACCCCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 795
DB 261 ThrValLeuThrGlyGlnArgProProThrGlnGlyProGlyProSerSerAsnIleThr 280
QY 796 TCTGGATCAGACCCCACTAGTCTTAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 855
DB 281 SerGlySerAspProThrGluSerAsnSerThrThrLysMetGlyAlaLysLeuPheSer 300
QY 856 CTCATCCAGGAGGCTTTTCAAGCTCTTAACTCCACGACTCCAGAGGCTACCTCTCTTGT 915
DB 301 LeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThrSerSerCys 320
QY 916 TGGCTATCTTACGCTTCGGGCCACCTTACTATGAAGGAATGGCTAGAGAGGGGAAATTC 975
DB 321 TrpLeuCysLeuAlaLeuGlyProProTrpTrpGlyGlyMetAlaArgArgGlyLysPhe 340
QY 976 AATGTGCAAAAGAACATAGAGACCAATGCACATGGGGATCCCAAAATAAGCTTACCCCTT 1035
DB 341 AsnValThrLysGlyHisArgAspProCysThrTrpGlySerGlnAsnLysLeuThrLeu 360
QY 1036 ACTGAGGTTCTGAAAGAGGCACCTGCATAGGAAAGGTTCCCCCATCCACCAACACCTT 1095
DB 361 ThrGluValPheGlyLysGlyThrCysIleGlyLysValProProSerHisGlnHisLeu 380
QY 1096 TGTAACACACACTGAAGCCTTTAATCAAACTCTCAGAGTCAATATCTGGTACCTGGTTAT 1155
DB 381 CysAsnHisThrGluAlaPheAsnArgThrSerGluSerGlnTrpLeuValProGlyTrp 400
QY 1156 GACAGGTGGTGGGCATGTAACTACTGGAATAACCCCTTGTGTTCACCTCGTGTTTTAA 1215
DB 401 AspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeuValPheAsn 420
QY 1216 CAAACTAAAGATTTTGCATTATGTCCTCAAAATTTGTTCCCGAGTGTATTACTATCCCGAA 1275
DB 421 GlnThrLysAspPheCysIleMetValGlnIleValProArgValTrpTrpTrpProGlu 440
QY 1276 AAAGCAATCCTTGATGAATATGACTACAGAATCATCGACAAAAGAGAGAACCCATATCT 1335
DB 441 LysAlaIleLeuAspGluTrpAspTrpArgAsnHisArgGlnLysArgGluProLysSer 460
QY 1336 CTGACACTTGTGTGATGCTCGGACTTCGGAGTGGCAGCAGGTGTAGGAAACAGGACAGCT 1395
DB 461 LeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGlyValGlyThrGlyThrAla 480
QY 1396 GCCTGTGTCAGGACCAACAGCAGCTAGAAACAGAGACTTAGTAACTACCTACGAAATGTA 1455
DB 481 AlaLeuValThrGlyProGlnGlnLeuGluThrGlyLeuSerAsnLeuHisArgIleVal 500
QY 1456 ACAGAGATCTCCAGGCCCTAGAAAATCTCTCAGTACCTCGGAGGAATCCCTAACCTCC 1515
DB 501 ThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluGluSerLeuThrSer 520
QY 1516 TTATCTGAAGTAGTCTACAGAATAGAGAGGGTTAGATTTTATTATTCTTAAAGAAGGA 1575
DB 521 LeuSerGluValValLeuGlnAsnArgGlyLeuAspLeuLeuPheLeuLysGluGly 540
QY 1576 GGATTATGTGTAGCCTTGAAGGAGGAATGCTGTTTTTATGTGGATCATTCAGGGCCATC 1635
DB 541 GlyLeuCysValAlaLeuLysGluCysCysPheTrpValAspHisSerGlyAlaIle 560
```

```
QY 1636 AGAGACTCCATGAACAGCTTAGAGAAAGGTTGGAGAGCGTCGAGGGAAGAAAGAACT 1695
Db 561 ArgAspSerMetSerLysLeuArgGluArgLeuGluLysArgArgGluLysGluThr 580
QY 1696 ACTCAAGGGTGTGGAGGATGTTCAACAGGCTCTTTGGTGGCTACCTACTTTCT 1755
Db 581 ThrGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpLeuAlaThrLeuLeuSer 600
QY 1756 GCTTTAACAGGACCTTAATAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1815
Db 601 AlaLeuThrGlyProLeuLeuValLeuLeuLeuLeuLeuLeuValGlyProCysIleIle 620
QY 1816 AACAGTTAATGCTCTCAITAGAGAACGAATAAGTCAGTCCAGATCATGGTACTTAGA 1875
Db 621 AsnLysLeuIleAlaPheIleArgGluArgIleSerAlaValGlnIleMetValLeuArg 640
QY 1876 CAACAGTACCAAGCCGCTAGCAGGGAAGCTGCGCG 1914
Db 641 GlnGlnTrpGlnSerProSerArgGluAlaGlyArg 653

RESULT 11
US-10-441-949-43
; Sequence 43, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodriago, Allen
; APPLICANT: Ross, Howard A.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-43

Alignment Scores:
Pred. No.: 1,49e-261 Length: 661
Score: 2879.50 Matches: 541
Percent Similarity: 88.3% Conservative: 40
Best Local Similarity: 82.2% Mismatches: 56
Query Match: 83.0% Indels: 21
DB: 4 Gaps: 6

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-43 (1-661)
QY 1 ATGCATCCCGAGTTAAACCGCGCCACTCCCGATTGGGGTGGAGGCGGAAAGCGAAAGACTG 60
Db 1 MethHisProThrLeuSerArgArgHisLeuProIleArgGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTTCATCGCGTGTCTTACTCTGTCAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATGGTATGCGCATAGGAGCAGCTGAACCTCCCATAAACCTTATCTCTCACTGG 180
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTACTGACTCCCGCAGCAGGTTAATATCAACAACTCAAGGGAGGCTCCCTTA 240
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 241 GSAACCTGGTGGCTGATCTATAGTTTGCTTCAGATCAGTTATTCCTAGTCTG----- 294
```

```
Db 81 GlyThrTrpTrpProGluLeuLysValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 295 ---ACCTCACCCAGATATCTCCATGCTCCTCAGGATTTTATGTTGGCCAGGACCA 351
Db 101 GlnAlaThrProProAspValLeuAlaTrpGlyPheTrpValCysProGlyProPro 120
QY 352 AATAATGGAAACATTCGGAAATCCAGAGATTTCTTTTGTAAACAATGAACCTGTGTA 411
Db 121 AsnAsnGluGluTrpCysGlyAsnProGlnAspPheCysLysGlnTrpSerCysVal 140
QY 412 ACCTCTAATGATGATATGGAATGCGCAACCTCTCAGCAGGATAGGGTAAAGTTTCT 471
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnAspArgValSerTrpSer 160
QY 472 TATGCTCAACACCTATACAGCTCTGGCAATTTAATTAC-----CTGACC 516
Db 161 PheValAsnAsnProThrSerTrpAsnGlnPheAsnTrpGlyHisGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGAACTGGAAGCCCAAGTCTCTCTCTTCAGAC 555
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysAsnSerLeuAsp 200
QY 556 CTAGATTACCTAAATAAGTTTCTCCTGAGAGAGAAACAAGAAAATATATCTCTAAATGG 615
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGluAsnIleGlnLysTrp 220
QY 616 GTAATGCTATGCTTCGGGAATGGTATATATGAGGCTCGGCTGAACAACACCGCTCC 675
Db 221 ValAsnGlyMetSerTrpGlyIleValTyrTrpGlySerGlyArgLysGlyLysTrp 240
QY 676 ATTCTAACTATTGCGCTCAAAATA---AACCAGCTGGAGCCCTCAATGGCTATAGGACCA 732
Db 241 ValLeuThrIleArgLeuArgGlyLeuThrGlnMetGluProProValAlaIleGlyPro 260
QY 733 AATACGGCTTTGACGGGTCAAAGACCCCAACCCAA-----GGACAGGACCA 780
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGluArgProSerProAsnPro 280
QY 781 TCCTCT---AACATACTTCTGGATCAGACCCCACTCAGTCTACAGCAGCACTAAATG 837
Db 281 SerAspTyrAsnThrTrpSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY 838 GGGGCAAACTTTTTCAGCTCATCCAGGAGCTTTTCAAGCTCTTAACCTCCACGACTCCA 897
Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY 898 GAGGCTACCTCTTCTTGTGCTATGCTTAGCTTCGGGCCCACTTACTATGAAGGAATG 957
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTrpGluGlyMet 340
QY 958 GCTAGAGAGGGAATTCATGTGACAAAAGACATAGACCCNATGCATGCGGATCC 1017
Db 341 AlaArgGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 1018 CAAAATAAGCTTACCTTACTGAGGTTCTCGAAAAGCGACCTGCATAGAAAGGTTCCC 1077
Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyArgValPro 380
QY 1078 CCAATCCCAACCAACCTTTGTAAACCACTGAAGCCCTTAAATCAAACTCTGAGAGTCAA 1137
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 1138 TATCTGCTACCTGTTATGACAGGTGGTGGCATGTAATCTGATTAACCCCTTGTT 1197
Db 401 TyrLeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 1198 TCCACCTTGTTGTTTAAACCAACTAAAGATTTTGCATTTATGGTCCAAATTTGTTCCCGCA 1257
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 1258 GTCTATTACTATCCCGAAAGCAATCCTTTGATGAATATGACTACAGAAATCATCGCAA 1317
```

Db 441 ValTyrTyrProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460
Qy 1318 AAGAGAACCCATATCTCTGACACTTCTCTGTGATGCTCGCACTTGGAGTGGCAGCAGGT 1377
Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
Qy 1378 GTAGAAACAGGAACAGCTGCTGTACAGGGACCACAGCAGCTAGAAACAGGACTTAGT 1437
Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
Qy 1438 AACCTACATCGAATTGTAAACAGAAGATCTCCAAGCCCTAGAAAAATCTGTCTAGTAACCTG 1497
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
Qy 1498 GAGGAATCCCTAACTCTTATCTGAAGTATGCTAGTCTTACAGATAGAAAGGGTTAGATTTA 1557
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
Qy 1558 TTATTCTTAAAGAGGAGGATTATGCTAGCCTTGAAGGAGGAATGCTGTTTTATGTG 1617
Db 541 LeuPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal 560
Qy 1618 GATCATTCAGGGGCCATCAGAGACTCCATGAACAGCTTAGAGAAAGGTTGGAGAAGCGT 1677
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArg 580
Qy 1678 CGAAGGGAAGGAACACTACTCAAGGTGGTTTGAAGGATGTTCAACAGGTCTCTTTGG 1737
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
Qy 1738 TTGGCTACCTTCTTCTGCTTAAACAGCACCTTAAATAGTCCTCTCTCTTACTCACA 1797
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThr 620
Qy 1798 GTTGGCCATGTATTATTAAACAAGTTAATTCGCTTTCATTAGAGAACGTAAGTGCAGTC 1857
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
Qy 1858 CAGATCATGGTACTTAGACAACAGTACCAAGCCCGTCTAGCAGGGAAGCTGGC 1911
Db 641 GlnIleMetValLeuArgGlnGlnTrpGlnGlyLeuLeuSerGlnGlyGluGly 658

RESULT 12

US-10-441-949-16
; Sequence 16, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; PRIORITY FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-16

Alignment Scores:
Pred. No.: 4.4e-261 Length: 660
Score: 2874.50 Matches: 539
Percent Similarity: 88.8% Conservative: 38
Best Local Similarity: 82.9% Mismatches: 52

Query Match: 82.9% Indels: 21
DB: 4 Gaps: 6
US-10-723-552-3_copy_5620_7533 (1-1914) x US-10-441-949-16 (1-660)
Qy 1 ATGCATCCCAAGTTAAACCGCGCCACCTCCGATTCGGGTGGAAGCCGAAAGACTG 60
Db 1 MethisProthrLeuSerArgArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
Qy 61 AAAATCCCTTAAAGCTTCGCTCCATCGCGTGGTTCTTACTCTGTCAATTAACCTCTCAG 120
Db 21 LysileProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
Qy 121 ACTAATGCGTATGCCATAGGACAGACCTCAACTCCCAATAAACCTTATCTCTCACTGG 180
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
Qy 181 TTAATTACTGATCCGCGCACAGGTATTAAATATCAACAACACTCAAGGGAGGCTCTTTTA 240
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
Qy 241 GGAACCTGGTGGCTGATCTATACGTTTGCCTCAGATCAGTATTCTTAGTCTG----- 294
Db 81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
Qy 295 ---ACCTCACCCAGATATCTCCATGCTCACGATTTTATGTTTGGCCAGGACCACCA 351
Db 101 GlnAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
Qy 352 AATAATGAAACAAATTCGGAAATCCCAGAGATTTCTTTTGTAAACAATGGAACCTGTGTA 411
Db 121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
Qy 412 ACCTCTAATGATGATNTTGGAAATGGCCCAACCTCTCAGCAGGATAGGTTTCTTCT 471
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpValSerGlnGlnAspArgValSerTyrSer 160
Qy 472 TATGTCACACCTATACCAGCTCTGGACAAATTTAATTAC-----CTGACC 516
Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
Qy 517 TGG-----ATTAGAACTCGAAGCCCCCAAGTGTCTCTCTTCAGAC 555
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
Qy 556 CTAGATTACTTAAATAAGTTTCTACTGAGAAAGGAAACAAAGAAATAATCTTAAATGG 615
Db 201 LeuAspTyrLeuLysIleSerPheThrLysGlyLysGlnGluAsnIleGlnLysTrp 220
Qy 616 GTAAATGCGTATGCTTGGGGAATGTTATATGAGGCTCGGTGTAACAAACACAGGCTCC 675
Db 221 ValAsnGlyMetSerTrpGlyIleValTyrTyrGlyGlySerGlyArgLysLysGlySer 240
Qy 676 ATTCTAATCTTCGCTCAAAATA---AACAGCTGGAGCCTCCCAATCGCTATAGGACCA 732
Db 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProProValAlaIleGlyPro 260
Qy 733 AATACGGTCTTGACGGGTCAAGAGCCCCCAACCCNA-----GGACCAGGACCA 780
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGluGlnArgProSerProAsnPro 280
Qy 781 TCCTCT---AACATAACTTCTGGATCAGACCCCACTGAGTCTAACAGCAGCAGCACTAAATG 837
Db 281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
Qy 838 GGGGCAAACTTTTTCAGCTCATCCAGGAGCTTTTCAAGCTCTTAACTCCACGACTCCA 897
Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
Qy 898 GAGGCTACCTCTTCTTGTGGCTATGCTTACCTGGGCCCCACCTTACTATGTAAGAATG 957
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMet 340


```
QY 958 GCTAGAGAGCGAAATTCATGTGACAAAAGACATAGACACCAATGACATGGGATCC 1017
Db 341 AlaArgGlyGlyLeuPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 1018 CAAATAAGCTTACCTTACTGAGGTTCTCGAAAGGCACCTGATAGAAAGGTTCC 1077
Db 361 GlnAsnLysLeuThrLeuGluValSerGlyLysGlyThrCysIleGlyArgValPro 380
QY 1078 CCATCCCAACCAACCTTTGTAACACCACTGAGCCTTTAATCAAACTCTGAGAGTCAA 1137
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 1138 TATCTGTACTGTTATGACAGGTGGTGGCATGTAATACCTGATTAACCCCTGTGTT 1197
Db 401 TyrLeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 1198 TCCACCTTGGTTTAAACCAACTAAAGATTTTGCATTATGCTCCAAATGTTCCCGA 1257
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 1258 GTGTATTACTATCCGAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGCAA 1317
Db 441 ValTyrTyrTrpProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460
QY 1318 AAGAGAGAACCATATCTCTGACACTTGTCTGTGATGCTCGGACTTGGAGTGGCAGG 1377
Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
QY 1378 GTAGAACAGGAACAGCTGCCCTCGTCACGGGACACAGCAGCTAGAAACAGGACTTAGT 1437
Db 481 ValGlyThrGlyThrAlaAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
QY 1438 AACCTACATCGAATTTGTAACAGAGATCTCCAAGCCCTAGAAAAATCTGTCAAGTAACTG 1497
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 1498 GAGGAATCCCTAACTCTTATCTGAAGTAGTCTTACAGATAGAGAGGTTAGATTTA 1557
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 1558 TTATTTCTAAAGAGAGGATATGTTAGCCTTGAAGGAGGAATCTGTTTTTATGTG 1617
Db 541 LeuPheLeuLysGluGlyGlyLeuCysValAlaLeuLysGluGluCysPheTyrVal 560
QY 1618 GATCATTCAGGGGCCATCAGAGACTCCATGAACCAAGCTTAGAGAAAGGTTGGAGAAGCGT 1677
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 580
QY 1678 CGAAGGAAAGGAAACTACTCAAGGGTGGTTGAGGGATGGTTCAACAGGTCTCTTTGG 1737
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 1738 TTGGCTACCTTACTTTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTGTTACTACA 1797
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuThr 620
QY 1798 GTTGGGCGATGTATTATTAAACAGTAAATTGCTTCAATGAGAACCAATAAGTGCAGTC 1857
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
QY 1858 CAGATCATGTTACTTAGACACAGTACCAA 1887
Db 641 GlnIleMetValLeuArgGlnGlnTyrGln 650
```

RESULT 13

US-10-441-949-8

; Sequence 8, Application US/10441949

; Publication No. US20040116684A1

; GENERAL INFORMATION:

; APPLICANT: Rodrigo, Allen

; APPLICANT: Ross, Howard A.

; APPLICANT: Mullins, James I.

; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES

```
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-8
```

Alignment Scores:

Pred. No.:	8,43e-261	Length:	660
Score:	2871.50	Matches:	539
Percent Similarity:	88.6%	Conservative:	37
Best Local Similarity:	82.9%	Mismatches:	53
Query Match:	82.8%	Indels:	21
DB:	4	Gaps:	6

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-8 (1-660)

```
QY 1 ATGCATCCCAAGTTAAACCGCGCCACTCCCGATTCGGGGTGGAAAGCCGAAAGACTG 60
Db 1 MethisProthrLeuSerArgHisLeuProIleArgGlyGlyLysProLysArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCTCCATCGCTGGTTCCTTACTCTGTCAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATGTTATGGCATAGGAGACAGCTCCATCAACTCCATAAACCTTATCTCTCACCTGG 180
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAATCACTCCGSCACAGGTATTAAATATCAACAACACTCAAGGGAGGCTCTTTA 240
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyGluAlaProLeu 80
QY 241 GGAACCTGTGTGGCTGTATCTATATAGTTTGCCTCAGATCAGTATTTCCTAGTCTG----- 294
Db 81 GlyThrTrpTrpProGluLeuTyrValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 295 ---ACCTCACCCAGATATCTCCATGCTCAGGATTTTATGTTGCCAGGACCAACA 351
Db 101 GlnAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
QY 352 AATAATGGAACCAATTCGGGAATCCAGAGATTTCTTTTGTAAACAATCGAACTGTGTA 411
Db 121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 412 ACCTCTAATCATGGATATTGGAATGCGCAACTCTCAGCAGGATAGGGTAAGTTTTCT 471
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY 472 TATGTCAACACCTATACAGCTCTGGACAATTTAATTAC-----CTGACC 516
Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGACTGGAGCCCAAGTCTCTCTCTCTCTGAC 555
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 556 CTAGATTACCTAAAAATAAGTTTCACTGAGAAAGGAAAAACAAGAAATATATCTTAAATGG 615
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyGlnGlnAsnIleGlnLysTrp 220
QY 616 GTAATATGGTATGCTTTGGGAATGGTATATTATGGAGGCTCGGTTAAACAACCGAGTCC 675
```


Db 101 GlnAlaThrProProAspValLeuArgAlaTyrGlyPheTyrValCysProGlyProPro 120
QY 352 AATAATGGAACAACTTGGGAAATCCAGAGATTCTTTTCTAACAATGAACTGTGTA 411
Db 121 AsnAsnGluGluTyrCysGlyAsnProGlnAspPheCysLysGlnTrpSerCysIle 140
QY 412 ACCTTAATGATGATATGGAATGGCCAACTCTCAGCAGGATAGGGTAAGTTTCT 471
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTyrSer 160
QY 472 TATGTCAACACTATACAGCTCTGGACAATTTAATTAC-----CTGACC 516
Db 161 PheValAsnAsnProThrSerTyrAsnGlnPheAsnTyrGlyHisGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGAAGCTGGAAGCCCAAGTGTCTCTCTCCCTCAGAC 555
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 556 CTAGATTACTTAAATAAGTTTCACTGAGAAAGGAAACCAAGAAATATATCTTAAATGG 615
Db 201 LeuAspTyrLeuLysIleSerPheThrGluLysGlyLysGlnGlnAsnIleGlnLysTrp 220
QY 616 GTAAATGATGTTCTGGGAATGGTATATTTATGAGGCTCGGTAAACAACAGGCTCC 675
Db 221 ValAsnGlyIleSerTyrGlyIleValTyrTyrGlyGlySerGlyArgLysLysGlySer 240
QY 676 ATTCTAACTATTCGCCCTCAAAATA--AACCAGCTGGAGCTCCCAATGGCTATAGGACCA 732
Db 241 ValLeuThrIleArgLysGluValGlnMetGluProProValAlaIleGlyPro 260
QY 733 AATACGGCTTTGACGGGTCAAGAGACCCCAACCCAA-----GGACCAAGGACCA 780
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGlnArgProSerProAsnPro 280
QY 781 TCCTCT---AACATACTCTGGATCAGACCCCACTGAGTCTTAACAGCAGCACTAAATG 837
Db 281 SerAspTyrAsnThrThrSerGlySerValProThrGluProAsnIleThrIleLysThr 300
QY 838 GGGGCAAACTTTTACGCTCATCCAGGAGCTTTCAAGCTCTTAAGTCTTAAGTCCAGCACTCA 897
Db 301 GlyAlaLysLeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrPro 320
QY 898 GAGGCTACCTCTCTCTGTTGGCTATGCTAGCTTCGGGCCCCACCTTACTATGAGGAATG 957
Db 321 GluAlaThrSerSerCysTrpLeuCysLeuAlaSerGlyProProTyrTyrGluGlyMet 340
QY 958 GCTAGAAGAGGAAATTCATGTGACAAAAGAACATAGACCAATGCATGGGGATCC 1017
Db 341 AlaArgGlyGlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySer 360
QY 1018 CAAAATAAGCTTACCTTACTAGGTTTCTGGAAAGGCCCTGCATAGGAAGGTTCCC 1077
Db 361 GlnAsnLysLeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyMetValPro 380
QY 1078 CCATCCCAACCAACCTTTGTAAACACACACTGAAGCTTTAATCAAACTCTGAGAGTCAA 1137
Db 381 ProSerHisGlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGln 400
QY 1138 TATCTGTACTCTGTTATGACAGTGTGGCATGTAAATCTAGGATTAACCCCTTGTTGTT 1197
Db 401 TyrLeuValProGlyTyrAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysVal 420
QY 1198 TCACCTTGTTTGTAAACCAACTTAAAGATTTTGCATTATGCTCCAAATGTTCCCGA 1257
Db 421 SerThrLeuValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArg 440
QY 1258 GTGTATTACTATCCCGAAAAAGCAATCTTGTATGAATATGACTACAGAAATCATCGCAA 1317
Db 441 ValTyrTyrTrpProGluLysAlaValLeuAspGluTyrAspTyrArgTyrAsnArgPro 460
QY 1318 AAGAGAAACCAATCTCTGACACTTGTGTGATGCTCGGACTTGGAGTGGCAGCAGGT 1377

Db 461 LysArgGluProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaAlaGly 480
QY 1378 GTAGAAACAGCAACAGCTGCCCTGTCACGGGACCACAGCAGCTAGAAACAGGACTTACT 1437
Db 481 ValGlyThrGlyThrAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSer 500
QY 1438 AACCTACATCGAATTGTAAACAGAGACTCTCCAAGCCCTAGAAAAATCTGTCAAGTAACTG 1497
Db 501 AsnLeuHisArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeu 520
QY 1498 GAGAAATCCCTAACCTCTTATCTGAAGTAGTCTCAGATAAGAGAGGTTAGATTTA 1557
Db 521 GluGluSerLeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeu 540
QY 1558 TTATTTCTAAAAGAGGAGGATTATGTAGCTTGAAGAGGAGGAAATGCTGTTTTTATGTCG 1617
Db 541 LeuPheLeuLysGluGlyLysCysValAlaLeuLysGluGluCysPheTyrVal 560
QY 1618 GATCATTTCAGGGCCATCAGAGACTCCATGAACAAAGCTTAGAGAAAGGTTGAGAAAGCGT 1677
Db 561 AspHisSerGlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArg 580
QY 1678 CGAAGGAAAGGAACTACTCAAGGCTGTTGAGGAGGATGTTCAACAGCTCTCTTTGG 1737
Db 581 ArgArgGluArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrp 600
QY 1738 TTGGCTACCTTCTCTGCTTTAAACAGGACCTTAATAGTCTCTCTCTGTTACTCACA 1797
Db 601 MetThrThrLeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuThr 620
QY 1798 GTTGGGCCATGTATTATTAACTTAATTTGCTTCAATTAGAGAACGAATAAGTGCAGTC 1857
Db 621 ValGlyProCysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaVal 640
QY 1858 CAGATCATGTACTTAGACAAACAGTACCAA 1887
Db 641 GlnIleMetValLeuArgGlnGlnTyrGln 650

RESULT 15
US-10-441-949-12
; Sequence 12, Application US/10441949
; Publication No. US20040116684A1
; GENERAL INFORMATION:
; APPLICANT: Rodrigo, Allen
; APPLICANT: Ross, Howard A.
; APPLICANT: Mullins, James I.
; TITLE OF INVENTION: ANCESTRAL VIRUSES AND VACCINES
; FILE REFERENCE: 08987-012001
; CURRENT APPLICATION NUMBER: US/10/441,949
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially generated peptide
US-10-441-949-12

Alignment Scores:
Pred. No.: 7,24e-259 Length: 678
Score: 2851.00 Matches: 539
Percent Similarity: 86.5% Conservative: 38
Best Local Similarity: 80.8% Mismatches: 52
Query Match: 82.2% Indels: 38
DB: 4 Gaps: 7

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-441-949-12 (1-678)

QY 1 ATGCATCCACGCTTAACCGCGCCACCTCCCGATTTCGGGGTGGAAACCGGAAAGACTG 60
Db 1 MethisProThrLeuSerArgHisIleuProIleArgGlyGlyProIleArgLeu 20
QY 61 AAAATCCCTTAAGCTTCGCGCTCCATCGCGTGGTTCTTACTCTGTCAATAACCTCTCAG 120
Db 21 LysIleProLeuSerPheAlaSerIleAlaTrpPheLeuThrLeuSerIleThrProGln 40
QY 121 ACTAATGGTATGCGCATAGGAGACAGCTGAACTCCGATAAACCTTATCTCTCACCTGG 180
Db 41 ValAsnGlyLysArgLeuValAspSerProAsnSerHisLysProLeuSerLeuThrTrp 60
QY 181 TTAATTAAGTCTCGCGCACAGGTATTAATATCAACACACTCAAGGGAGGCTCCTTTA 240
Db 61 LeuLeuThrAspSerGlyThrGlyIleAsnIleAsnSerThrGlnGlyAlaProIleu 80
QY 241 GGAACCTGGTGGCTGATCTATACGTTTGGCTTCAGATCAGTATTATCTCTAGTCTG 294
Db 81 GlyThrTrpTrpProGluLeuTyValCysLeuArgSerValIleProGlyLeuAsnAsp 100
QY 295 ---ACCTCACCCCGAGATATCTCATGCTCAGCGATTTTATGTTTGGCCAGGACCA 351
Db 101 GlnAlaThrProProAspValLeuArgAlaTyArgPheTyValCysProGlyProPro 120
QY 352 AATAATGGAAACATTCGGGAATCCAGAGATTCTTTCTTAAACATGGAATGTA 411
Db 121 AsnAsnGluGluTyCysGlyAsnProGlnAspPhePheCysLysGlnTrpSerCysVal 140
QY 412 ACCTCTAATGATGATATTCGGAATGGCCAACTCTCAGCAGGATAGGGTAAGTTTCT 471
Db 141 ThrSerAsnAspGlyAsnTrpLysTrpProValSerGlnGlnAspArgValSerTySer 160
QY 472 TATGTCAACACTATACGAGCTCTGGACAATTTAATTAC-----CTGACC 516
Db 161 PheValAsnAsnProThrSerTyAsnGlnPheAsnTyArgHisGlyArgTrpLysAsp 180
QY 517 TGG-----ATTAGAAGTGGAGCCCAAGTCTCTCTCTCAGAC 555
Db 181 TrpGlnGlnArgValGlnLysAspValArgAsnLysGlnIleSerCysHisSerLeuAsp 200
QY 556 CTAGATTACCTAAATAAGTTTCACTCAGAAAGGAAACAAAGAAATATCTTAAATGG 615
Db 201 LeuAspTyLeuLysIleSerPheThrGlnLysGlyLysGlnLysGlnIleGlnLysTrp 220
QY 616 GTAATGTATGCTCTGGGGAATGTATATATGAGGCTCGGTAAACAACACGAGCTCC 675
Db 221 ValAsnGlyMetSerTrpGlyLeValTyTyArgGlySerGlyArgLysLysGlySer 240
QY 676 ATTCTAAGTATTCGCTCAAAATA---AACAGCTGGAGCCTCCCAATGGCTATAGGACCA 732
Db 241 ValLeuThrIleArgLeuArgIleGluThrGlnMetGluProProValAlaIleGlyPro 260
QY 733 AATACGGTCTGACGGGTCAAGAGACCCCAACCCCAAGGACCA----- 774
Db 261 AsnLysGlyLeuAlaGluGlnGlyProProIleGlnGluProProHisAsnLeuProVal 280
QY 775 -----GGACCACTCTCT-----AACATAACTTCTCGA 801
Db 281 ProGlnArgProSerProAsnProAspIleThrGlnSerAspTyArgAsnThrThrSerGly 300
QY 802 TCAGACCCCACT-----GAGTCTAACACACAGCTAAATAATGGGGCAAAA 846
Db 301 SerValProThrAsnThrProArgAsnGluProAsnIleThrIleLysThrGlyAlaLys 320
QY 847 CTTTTAGCTCATCAGGAGCTTTTCAAGCTCTTAACTCTCAACGACTCCAGAGGCTACC 906
Db 321 LeuPheSerLeuIleGlnGlyAlaPheGlnAlaLeuAsnSerThrThrProGluAlaThr 340
QY 907 TCTTCTGTGTGCTATGCTTACGTTTCGGGCCCACTTATCTATGAAGGAATGGCTAGAAGA 966
Db 341 SerSerCysTrpLeuCysLeuAlaSerGlyProProTyTyArgGlyMetAlaArgGly 360

QY 967 GGGAAATTCATGTGACAAAGAAACATAGAGACCAATGCATGGGGATCCAAAATAAG 1026
Db 361 GlyLysPheAsnValThrLysGluHisArgAspGlnCysThrTrpGlySerGlnAsnLys 380
QY 1027 CTTACCTTACTGAGGTTTCTGGAAAAGGACCTGCGATAGGAAAAGTTCCCCCATCCAC 1086
Db 381 LeuThrLeuThrGluValSerGlyLysGlyThrCysIleGlyArgValProProSerHis 400
QY 1087 CAACACCTTTGTAAACACACACTGAAGCCTTTAATCAAACCTCTGAGAGTCAATATCTGTA 1146
Db 401 GlnHisLeuCysAsnHisThrGluAlaPheAsnArgThrSerGluSerGlnTyLeuVal 420
QY 1147 CCTGGTTATGACAGCTGGTGGCCATGTAATACTCGATTAAACCCCTGTGTTCACACCTTG 1206
Db 421 ProGlyTyArgAspArgTrpTrpAlaCysAsnThrGlyLeuThrProCysValSerThrLeu 440
QY 1207 GTTTTTAACCAAACTAAGATTTTTGCAATTATGTCCAAATTTGTTCCCGAGTGTATTAC 1266
Db 441 ValPheAsnGlnThrLysAspPheCysValMetValGlnIleValProArgValTyTy 460
QY 1267 TATCCCGAAAGCAATCCTTGATGAATATGACTACAGAAATCATCGACAAAGAGAGAA 1326
Db 461 TyrProGluLysAlaValLeuAspGluTyArgTyArgTyArgAsnArgProLysArgGlu 480
QY 1327 CCCATATCTCTGACACTTGTCTGATGCTCGACTTGGAGTGGCAGAGGTGTAGGAAACA 1386
Db 481 ProIleSerLeuThrLeuAlaValMetLeuGlyLeuGlyValAlaIleGlyValGlyThr 500
QY 1387 GGAACAGCTGCCCTGGTCCGCGGACCAACAGCAGCTAGAAACAGCACTTAGTAACTACAT 1446
Db 501 GlyThrAlaLeuIleThrGlyProGlnGlnLeuGluLysGlyLeuSerAsnLeuHis 520
QY 1447 CGAATTGTAAACAGAAATCTCCAGCCCTAGAAAATCTGTCAGTAACTCGAGGAATCC 1506
Db 521 ArgIleValThrGluAspLeuGlnAlaLeuGluLysSerValSerAsnLeuGluSer 540
QY 1507 CTAACCTCTCTATCTGAAGTAGTCTCTACAGAAATAGAAGAGGTGTAGATTATTATTCTA 1566
Db 541 LeuThrSerLeuSerGluValValLeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeu 560
QY 1567 AAAGAAGAGGATTTATGTGTAGCCTTGAAGAGGAATGCTCTTTTATGTGGATCATTTCA 1626
Db 561 LysGluGlyGlyLeuCysValAlaLeuLysGluCysCysPheTyValAspHisSer 580
QY 1627 GGGGCCATCAGAGACTCCATGAACAAGCTTAGAAAAGTTTGGAGAAGCGTCGAGGGA 1686
Db 581 GlyAlaIleArgAspSerMetSerLysLeuArgGluArgLeuGluArgArgArgGlu 600
QY 1687 AAGGAAACTACTCAAGGTGGTTTGGAGGATGGTTTCAACAGGTCTCTTTGGTGGCTACC 1746
Db 601 ArgGluAlaAspGlnGlyTrpPheGluGlyTrpPheAsnArgSerProTrpMetThrThr 620
QY 1747 CTACTTTCTGCTTTAACAGGACCCCTTAATAGTCTCTCTCTGTTACTCAGAGTTGGGCA 1806
Db 621 LeuLeuSerAlaLeuThrGlyProLeuValValLeuLeuLeuLeuLeuValGlyPro 640
QY 1807 TGTATTATTAAACAGTTAATTCCTTCAATTAGACAGCAAGTAAGTCAGTCCAGATCATG 1866
Db 641 CysLeuIleAsnArgPheValAlaPheValArgGluArgValSerAlaValGlnIleMet 660
QY 1867 GTACTTGAACAACAGTACCAA 1887
Db 661 ValLeuArgGlnTyGln 667

Search completed: February 14, 2006, 17:19:42

Job time : 288.633 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 14, 2006, 16:19:01 ; Search time 1.67839 Seconds
(without alignments)
2992.933 Million cell updates/sec

Title: US-10-723-552-3_COPY_5620_7533

Perfect score: 3468

Sequence: 1 ATGCATCCAGGTTAAACCG.....CTACGAGGAGAGCTGGCCGC 1914

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 194028

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10723552/runat 14022006 125151 13368/app query.fasta_1
-DB=Published Applications AA New -OFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biolum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -HOST=abs02p
-USRR=US10723552 @CGN 1.1.17 @runat 14022006 125151 13368 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -JONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELETE=7

Database : Published Applications AA New:

- 1: /cgn2_6/ptodata/1/pubpaa/US08 NEW PUB pep:**
- 2: /cgn2_6/ptodata/1/pubpaa/US06 NEW PUB pep:**
- 3: /cgn2_6/ptodata/1/pubpaa/US07 NEW PUB pep:**
- 4: /cgn2_6/ptodata/1/pubpaa/US05 NEW PUB pep:**
- 5: /cgn2_6/ptodata/1/pubpaa/US10 NEW PUB pep:**
- 6: /cgn2_6/ptodata/1/pubpaa/US11 NEW PUB pep:**
- 7: /cgn2_6/ptodata/1/pubpaa/US10 NEW PUB pep:**
- 8: /cgn2_6/ptodata/1/pubpaa/US06 NEW PUB pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1267.5	36.5	654	US-10-510-947-6	Sequence 6, Appli
2	413.5	11.9	540	US-10-821-234-1395	Sequence 1395, Ap
3	250	7.2	229	US-11-014-629-1	Sequence 1, Appli
4	170.5	4.9	1189	US-10-821-234-1209	Sequence 1209, Ap
5	127.5	3.7	676	US-10-510-947-1	Sequence 1, Appli
6	118	3.4	490	US-10-131-826A-310	Sequence 310, App
7	101.5	2.9	1225	US-11-102-476-2	Sequence 2, Appli
8	100	2.9	566	US-11-033-039-1244	Sequence 1244, Ap
9	97.5	2.8	748	US-11-098-686-10863	Sequence 10863, A

10	97	2.8	410	6	US-10-821-234-1180	Sequence 1180, Ap
11	96.5	2.8	801	7	US-11-014-842A-35	Sequence 35, Appl
12	96.5	2.8	869	7	US-11-014-842A-27	Sequence 27, Appl
13	96	2.8	742	7	US-11-175-690-528	Sequence 528, App
14	96	2.8	742	7	US-11-175-690-534	Sequence 534, App
15	96	2.8	742	7	US-11-175-690-546	Sequence 546, App
16	95	2.7	1985	6	US-11-095-083-7	Sequence 7, Appli
17	93.5	2.7	1255	7	US-11-022-562-235	Sequence 235, App
18	93.5	2.7	1255	7	US-11-052-554A-265	Sequence 265, App
19	93.5	2.7	1255	7	US-11-052-554A-266	Sequence 266, App
20	93.5	2.7	1255	7	US-11-052-554A-267	Sequence 267, App
21	93.5	2.7	1255	7	US-11-052-554A-268	Sequence 268, App
22	93.5	2.7	1255	7	US-11-052-554A-269	Sequence 269, App
23	93.5	2.7	1255	7	US-11-052-554A-270	Sequence 270, App
24	93.5	2.7	1255	7	US-11-052-554A-271	Sequence 271, App
25	93.5	2.7	1255	7	US-11-052-554A-272	Sequence 272, App
26	93.5	2.7	1255	7	US-11-052-554A-274	Sequence 274, App
27	93.5	2.7	1255	7	US-11-052-554A-275	Sequence 275, App
28	93.5	2.7	1279	6	US-10-957-880-3	Sequence 3, Appli
29	92.5	2.7	595	7	US-11-102-476-33	Sequence 33, Appl
30	92.5	2.7	3011	6	US-10-985-205-3	Sequence 3, Appli
31	92.5	2.7	8746	7	US-11-098-686-10232	Sequence 10232, A
32	92	2.7	515	7	US-11-052-554A-255	Sequence 255, App
33	92	2.7	1237	7	US-11-052-554A-95	Sequence 95, Appl
34	92	2.7	3194	7	US-11-052-554A-90	Sequence 90, Appl
35	91.5	2.6	922	7	US-11-115-086-9	Sequence 9, Appli
36	90.5	2.6	398	6	US-10-517-939-210	Sequence 210, App
37	90.5	2.6	757	7	US-11-052-554A-378	Sequence 378, App
38	90.5	2.6	899	6	US-10-453-372-528	Sequence 528, App
39	90.5	2.6	5935	6	US-10-995-561-776	Sequence 776, App
40	90	2.6	567	7	US-11-033-039-1279	Sequence 1279, Ap
41	89.5	2.6	536	7	US-11-076-074-2	Sequence 2, Appli
42	89.5	2.6	898	6	US-10-453-372-522	Sequence 522, App
43	89.5	2.6	899	6	US-10-453-372-506	Sequence 506, App
44	89.5	2.6	899	6	US-10-453-372-524	Sequence 524, App
45	89.5	2.6	899	6	US-10-453-372-526	Sequence 526, App

ALIGNMENTS

RESULT 1
US-10-510-947-6
; Sequence 6, Application US/10510947
; Publication No. US2005055123A1
; GENERAL INFORMATION:
; APPLICANT: The Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Medina, Maria Fe C.
; APPLICANT: Kobinger, Gary
; TITLE OF INVENTION: Chimeric Ebola Virus Envelopes and Uses Therefor
; FILE REFERENCE: UPN-02811PCT
; CURRENT APPLICATION NUMBER: US/10/510,947
; CURRENT FILING DATE: 2004-10-28
; PRIOR APPLICATION NUMBER: US 60/376,480
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/385,704
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/427,752
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Murine leukemia virus
US-10-510-947-6
Alignment Scores:
Pred. No.: 4.15e-111 Length: 654
Score: 1267.50 Matches: 277
Percent Similarity: 57.0% Conservatives: 91
Best Local Similarity: 42.9% Mismatches: 189
Query Match: 36.5% Indels: 89

[illegible]

RESULT 3
US-11-014-629-1

```

US Air Force 025-1
Sequence 1, Application US/11014529
Publication NO. US20050244376A1
GENERAL INFORMATION:
APPLICANT: Hall, Frederick L.
TITLE OF INVENTION: TARGETED GENE 1
FILE REFERENCE: 30863-704.302
CURRENT APPLICATION NUMBER: US/11/
CURRENT FILING DATE: 2004-12-15

```

;; PRIORITY APPLICATION NUMBER: US 08/837,223
;; PRIORITY FILING DATE: 1997-04-10
;; PRIORITY APPLICATION NUMBER: US 09/904,923
;; PRIORITY FILING DATE: 2001-07-13
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: moloney murine leukemia virus
;; US-11-014-629-1

Alignment Scores:
Pred. No.: 1,05e-15 Length: 229
Score: 250.00 Matches: 70
Percent Similarity: 39.1% Conservative: 23
Best Local Similarity: 29.4% Mismatches: 75
Query Match: 7.2% Indels: 70
DB: 7 Gaps: 12

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-11-014-629-1 (1-229)

```
Qy 157 CATAAACCTTATCTCTCACCTGGTTAAATTAATGACTCCGGCACAGGATTAATTAATCAAC 216
Db 8 HisGlnValTyrAsnIleThrTIPgluValThr---AsnGlyAspArgGluThrValTrp 26
Qy 217 AACACTCAGGGAGGCTCCCTTTAGGAACCTGGTGGCTGATCTATAC-----GTTTGC 270
Db 27 AlathrSerGlyAsnHisProLeuTrpTrpTrpProAspLeuThrProAspLeuCys 46
Qy 271 CTC-----AGATCATGTTATTCCTAGT 291
Db 47 MetLeuAlaHisGlyProSerTyrTrpGlyLeuGluTyrGlnSerProPheSerSer 66
Qy 292 CTGACCTCACCCCA----- 306
Db 67 ProProGlyProProCysCysSerGlySerSerProGlyCysSerArgAspCysGlu 86
Qy 306 ----- 306
Db 87 GluProLeuThrSerLeuThrProArgCysAsnThrAlaTrpAsnArgLeuLysLeuAsp 106
Qy 307 GATATCTCCATGCT-----CACGATTTTATGTTGCCAGGACCA-----CCAAAT 354
Db 107 GlnThrThrHisLysSerAsnGluGlyPheTyrValCysProGlyProHisArgProArg 126
Qy 355 AATGAAACATTCGGGAAATCCAGAGATTTCTTTTAAACAATGGAACCTGTGTAAAC 414
Db 127 GluSerLysSerCysGlyGlyProAspSerPheTyrCysAlaTyrTrpGlyCysGluThr 146
Qy 415 TCTAATGATGATATTTGGAATGGCCAACTCTCAGCAGGATAGGTTAAGTTTCTTTAT 474
Db 147 ThrGlyArgAlaTyrTrpLys---ProSerSerSerTrpAsp-----PheIleThr 162
Qy 475 GTCAACACTATACCACTCTGGACAATTTAATACCTGACCTGGATAGACTGGAAC 534
Db 163 ValAsnAsnLeuThrSerAspGln-----AlaValGlnValCysLysAspAsn 179
Qy 535 CCCAAGTCTCTCCCTTCAGACCTAGATTACCTAAAATAAGTTTCACTGAGAAAGGAAA 594
Db 180 LysTrpCysAsnPro-----LeuValIleArgPheThrAspAlaGlyArg 194
Qy 595 CAAGAAATATCTTAAATGGGTAAATGGTATGTTGGGAAATGGTATATATTAGGAGC 654
Db 195 Arg-----ValThrSerTrpThrThrGlyHisTyrTrpGlyLeuArgLeuTyr---Val 211
Qy 655 TCGGTAACACACAGGCTCCATTCTACTATTTCGCTCAAAATAAACCCAGCTG 708
Db 212 SerGlyGlnAspProGlyLeuThrPheGlyIleArgLeuArgTyrGlnAsnLeu 229
```

RESULT 4

US-10-821-234-1209
; Sequence 1209, Application US/10821234

;; Publication No. US20050255114A1
;; GENERAL INFORMATION:
;; APPLICANT: Labat, Ivan
;; APPLICANT: Stache-Crain, Birgit
;; APPLICANT: Andarmani, Susan
;; APPLICANT: Tang, Y. Tom
;; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;; FILE REFERENCE: 821A
;; CURRENT APPLICATION NUMBER: US/10/821,234
;; CURRENT FILING DATE: 2004-04-07
;; PRIORITY APPLICATION NUMBER: US 60/462,047
;; PRIORITY FILING DATE: 2003-04-07
;; NUMBER OF SEQ ID NOS: 1704
;; SOFTWARE: pt_seq_genes Version 1.0
;; SEQ ID NO 1209
;; LENGTH: 1189
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(1189)
;; OTHER INFORMATION: xaa = any amino acid or nothing
;; US-10-821-234-1209

Alignment Scores:
Pred. No.: 4,92e-08 Length: 1189
Score: 170.50 Matches: 45
Percent Similarity: 48.1% Conservative: 17
Best Local Similarity: 34.9% Mismatches: 34
Query Match: 4.9% Indels: 33
DB: 6 Gaps: 4

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-821-234-1209 (1-1189)

```
Qy 1468 CAAGCCCTAGAAAA-----TCTGTAGTAACTGGAGGAATCCCTAAC----- 1512
Db 1075 GlnAlaLeuProLysProHisGluGlyValSerAsnProLys**ThrCysSerProIle 1094
Qy 1513 -----TCTTATCTGAAGTAGTCTTACAGATAGAGGGGTTAGATTATTATT 1563
Db 1095 ProLysAspSerLeuSerArgValThrLeuGlnAsnHisArgGlyLeuAspLeuThr 1114
Qy 1564 CTAAGAAGAGGATATGTGTAGCTTGAAGAGGAATGCTTTTATGTGGATCAT 1623
Db 1115 AlaGluLysGlyGlyLeuCysIlePheLeuGluGluCysCysPheTyrThrAsnGln 1134
Qy 1624 TCAGGGCCCATCAGACACTCCATGAACAAGCTTAGAGAAAGGTTGGAGAACGTCGAAG 1683
Db 1135 SerGlyLeuValGlnAspAlaAlaGlyArgIleAsnGluLysAlaSer----- 1150
Qy 1684 GAAAGGAACACTACTCAAGGTTGTTGAGGGATGGTTCAACAGTCTCTTTGGTTGGCT 1743
Db 1151 -----GlyArgValGlnTrpLeuThr 1157
Qy 1744 ACCCTACTTTCTGCTTTA-----ACAGGACCCCTTAATAGTC 1779
Db 1158 ProValIleProAspLeuTrpGluAlaGluAlaGlyGlySerArgGlyGlnGluLeu 1177
Qy 1780 CTCCTCTCTGTTACTCACAGTTGGGCCA 1806
Db 1178 ThrIleLeuAlaAsnThrValLysPro 1186
```

RESULT 5

US-10-510-947-1
; Sequence 1, Application US/10510947
; Publication No. US20050255123A1
;; GENERAL INFORMATION:
;; APPLICANT: The Trustees of the University of Pennsylvania
;; APPLICANT: Wilson, James M.
;; APPLICANT: Medina, Maria Fe C.
;; APPLICANT: Kobinger, Gary
;; TITLE OF INVENTION: Chimeric Ebola Virus Envelopes and Uses Therefor
;; FILE REFERENCE: UPN-02811pCT

;; CURRENT APPLICATION NUMBER: US/10/510,947
;; CURRENT FILING DATE: 2004-10-28
;; PRIOR APPLICATION NUMBER: US 60/376,480
;; PRIOR FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/385,704
;; PRIOR FILING DATE: 2002-06-04
;; PRIOR APPLICATION NUMBER: US 60/427,752
;; PRIOR FILING DATE: 2002-11-20
;; NUMBER OF SEQ ID NOS: 67
;; SOFTWARE: Patent in version 3.2
;; SEQ ID NO 1
;; LENGTH: 676
;; TYPE: PRT
;; ORGANISM: Ebola virus
US-10-510-947-1

Alignment Scores:

Pred. No.:	0.000455	Length:	676
Score:	127.50	Matches:	103
Percent Similarity:	36.1%	Conservative:	88
Best Local Similarity:	19.5%	Mismatches:	206
Query Match:	3.7%	Indels:	132
DB:	6	Gaps:	23

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-10-510-947-1 (1-676)

Qy	439	CCAACTCTCAGCAGGATAGGTAAGTTTCTTATGTCAACACC-----TATACC	489
Db	202	ProValAsnAlaThrGluAspProSerSerGlyTyrThrThrThrIleArgTyrGln	221
Qy	490	AGCTCTCGA-----CAATTAATTAACCTGACCTGG	519
Db	222	AlaThrGlyPheGlyThrAsnGluThrGluTyrLeuPheGluValAspAsnLeuThrTyr	241
Qy	520	ATTAGAACTGGAGCCCAAGTCTCTCCACAGCTAGATTACCTAAATAAATAGTTTC	579
Db	242	ValGlnLeuGluSer---ArgPheThrProGlnPheLeuGlnLeuAsnGluThrIle	260
Qy	580	ACTGAGAAAGGAAACCAAGAAATATC-----CTAAATGG-----	615
Db	261	TyrThrSerGlyLysAsnThrThrGlyLysLeuIleTyrLysValAsnProGlu	280
Qy	616	GTAATGGTATGCTTGGGAATGGTATATATAGGAGCTCGGGAACCAACAGGCTCC	675
Db	281	IleAsp---ThrThrIleGlyGluTyr-AlaPheTyrGlu-----ThrLysLysA	296
Qy	676	ATTCTAACTATCGCTCAAAATAAACACAGCTGGAGCTCCATGGCTATAGACCAAT	735
Db	296	sn-LeuThrArgLysIleAsnGluLeuSerPheThrValValSerAsnGlyAla	315
Qy	736	ACGGTCTTGACGGTCAAGACCCCAACCAAGGACGACCATCTCTAACAATAACT	795
Db	316	LysAsnIleSerGlyGlnSerProAlaArgThrSerSerAspProGlyThrAsnThr	335
Qy	796	TCGTGATCAGACCCCACTGATCTTAACAGCACGACTAAATAGGGGGCAAACTTTT	855
Db	336	ThrGluAspHisLysIleMetAlaSerGluAsnSerSerAlaMetValGlnValHis	355
Qy	856	CTCATCCAGGAGCTTTTCAAGCTCTTAATCCACGACTCAGAGGCTACTCTTCTGT	915
Db	356	-----GlnGlyArgGluAlaAlaValSerHisLeuThrThrLeuAlaThr-----	370
Qy	916	TGGCTATGCTTAGTCTCGGGCCCACTTACTATGAAGGAATGGCTAGAGGGAATTC	975
Db	371	-----IleSerThrSerPro-----GlnSer	377
Qy	976	AATGTGCAAAAGAACATAGACCAATGCACATGGGGATCCCAA---AATAAGCTTACC	1032
Db	378	LeuThrThrLysProGlyProAspAsnSerThrHisAsnThrProValTyrLysLeuAsp	397
Qy	1033	CTTACTAGAGTTTCTGGAAGGACCTGCATAGGAAGGTTCCCATCCCAACACAC	1092
Db	398	IleSerGluAlaThr-----GlnValGluGlnHisArgArg	410

Qy	1093	CTTTGTAAACCACTGAAGCCTTTAATCAAACTCTGAGAGTCAATATCTGTACTCGT	1152
Db	411	ThrAspAsnAspSerThrAlaSerAspThrProSerAlaThrThrAlaAlaGlyPro---	429
Qy	1153	TATCAGAGGTGGTGGGCATGTAATCTGATTAACCCCTTGTGTTTCCACCTGGTTT	1212
Db	430	-----ProLysAlaGluAsnThrAsnThr	437
Qy	1213	AACCAAAATAAGATTTTTCATTATGTCCTCAAAATGTTCCCGAGTGATTACTATCC	1272
Db	438	SerLysSerThrAspPhe-----LeuAspProAlaThrThrThrSerPro	452
Qy	1273	GAATAAACAATCTCTTGATGAATATGACTACAGAAATCATCGACAA-----	1317
Db	453	GlnAsnHisSerGluThrAlaGlyAsnAsnAsnThrHisHisGlnAspThrGlyGlu	472
Qy	1318	-----AAGAGAGAACCCTATCTCTGCACACTTGTGTGATGCTCGGACTT	1362
Db	473	SerAlaSerSerGlyLysLeuGlyLeuIleThrAsnThrIleAlaGlyValAlaGlyLeu	492
Qy	1362	-----	1362
Db	493	IleThrGlyArgArgThrArgArgGluAlaIleValAsnAlaGlnProLysCysAsn	512
Qy	1363	-----GGAGTGGCAGCAGGTGTAGGA-----	1383
Db	513	ProAsnLeuHisTyrTyrThrThrGlnAspGluGlyAlaIleGlyLeuAlaTyrPle	532
Qy	1384	-----ACAGAAACAGCTGCC-----CTGTCACGGGACACACAGCAGCTAGAAACA	1428
Db	533	ProTyrPheGlyProAlaAlaGluGlyIleTyrIleGluGlyLeuMetHisAsnGlnAsp	552
Qy	1429	GGACTT---AGTAACCTACATCGAATTTGTAACAGAGATCTCCAGCCCTAGAAAAATCT	1485
Db	553	GlyLeuIleCysGlyLeuArgGlnLeuAlaAsnGluThrThrGlnAlaLeuGlnLeuPhe	572
Qy	1486	GTCAAGTAACCTGGAGGAATCCCTAACTCTCTCAAGTAGTCTACAGATAGAGA	1545
Db	573	Leu-----ArgAlaThrThrGluLeuArgThrPheSerIleLeuAsnArgLys	588
Qy	1546	GGGTAGATTATTTCTTAAAGAGAGGAGGATATGTGTAGCTTTGAAGGAGGAATGC	1605
Db	589	AlaIleAspPheLeuLeuGlnArgTyrGlyThrCysHisIleLeuGlyProAspCys	608
Qy	1606	TGTTTATATCG---GATCATTGAGGGCCATCAGACACTCCATGAACAAG---CTTAGA	1659
Db	609	CysIleGluProHisAspTyrThrLysAsnIleThrAspLysIleAspGlnIleHis	628
Qy	1660	GAAAGGTGGAGAGCGTCGAAGGAAAGAACTACTCAAGGTGTTTGGGGATGG	1719
Db	629	AspPheValAspLysThrLeuProAspGlnGlyAspAsnAspAsnThrThrGlyTyr	648
Qy	1720	TTCAACAGGTCTCTTTGGTGGCTACCCCTACTTCTGCTTTAACAGGACCCCTTAATAGTC	1779
Db	649	-----ArgGlnTyrPleProAlaGlyIle---GlyValThrGlyValIleAla	664
Qy	1780	CTCCTCTCTTACTACAGTTT	1800
Db	665	ValleAlaLeuPheCysIle	671

RESULT 6

US-10-131-826A-310
; Sequence 310, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.

; ORGANISM: Leptospira interrogans
US-11-102-476-2

Alignment Scores:	
Pred. No.:	0.149
Score:	101.50
Percent Similarity:	35.7%
Best Local Similarity:	22.9%
Query Match:	2.9%
DB:	7
Length:	1225
Matches:	73
Conservative:	41
Mismatches:	121
Indels:	84
Gaps:	16

US-10-723-552-3 COPY 5620 7533 (1-1914) x US-11-102-476-2 (1-1225)

693	QY	CAAATAAACAAGCTGGAGCCTCC-----AATGGCTATAGGACCAAAATACGGTCTCTGAC	746
587	Db	GlnIleThrProAlaAlaSerLysAlaLysGlyLeuThrGluArgPheLysAlaThr	606
747	QY	GGGTCAAGACCCCAACCCAGGACGACCACTCTTAAATACATCTCTGGATCAGA	806
607	Db	GlyIlePheThrAspAsnSerAsnSerAspIle-ThrAsnGlnValThrTrpAsnSerSe	626
807	QY	CCCCACTGAG-----TCTAACAGCAGCACTAAAATGGG-----	840
626	Db	rAsnThrAspIleLeuThrValSerAsnThrAsnAlaLysArgGlyLeuGlySerThrLe	646
841	QY	-----GCAAAACTTTTTCAGCTCATCCAGGAGCTTTTCAAGCTCTTCAACTC	887
646	Db	uLysGlnGlyThrValLysValThrAlaSerMetGlyGlyIleGluAspSerValAspH	666
888	QY	CACGACTCCAGAGGTACTCTCTTGTGTGGCTATGCTAGCTTCGGGCCCACTTACTA	947
666	Db	eThrValThrGlnAlaThrLeuThrSerIleGluValSerProThrArgAlaSerIleAl	686
948	QY	TCAGCGAATGGCTAGACAGGAGAAATCAATGTGACA-----	984
686	Db	alysGlyMetThrGln-----LysPheThrAlaThrGlyIlePheThrAspHisSerLy	704
985	QY	AAAGAACATAGAGACCAATGCATGGGGATCCAAAATAAG-----CTTACCCCTT	1033
704	Db	sLysAsnIleThrGlnGlnValThrTrpLysSerSerLysAlaLeuSerMetLeuAs	724
1036	QY	-----ACTGAGGTTCTCGAAAAGGCACCTGCATAGGAAGGTTCCCCCACTCCA	1085
724	Db	nAlaProGlyGluGluGlyThrGlyLysAlaIleAlaValGlyLys-----Hi	740
1086	QY	CCACACACTTGTTAACCACTGAAGCCTTTAATCAAACTCTGAGAGTCATATCTGGT	1145
740	Db	sTyrTyrCysAsnLeuArgLysThrPheArgGluAsnArgTyrTyrArgTyrSerAr	760
1146	QY	ACCTGGTTAT---GACAGGTGGTGGCATGTATACT-----GGATTAAAC	1187
760	Db	gAsnSerTyrPheAsnSerAsnGlnSerCysLysAsnIleValLeuValLysGlyLeuTh	780
1188	QY	CCCTTGTGTTTCC---ACCTTGGTTTTT-----AACCAACTAAAGATTTTTGCATT	1238
780	Db	rGluLysPheSerAlaThrGlyIleTyrSerAspAsnSerSerLysAsp-----	796
1239	QY	GGTCCAAATTGTTCCCGAGTGTTACTATCCGAAAAGCAATCTTGTAGTAATGA	1295
797	Db	-----IleThrSerAlaValTh	802
1299	QY	CTACAGAAATCATCGACAAAGAGAGAACCCATATCTTGACACTTGTGTGATGCTCGG	1355
802	Db	rTrpHisSerSerAsnSerValAlaThrIleSerAsnThrLysGlyTyrGln-----	820
1359	QY	ACTTGGAGTGGCAGCGGTGTAGCAACAGGAACAGCTGCCCTGGTCACGGGACCACGA	1418
821	Db	---GlyGlnAlaHisGlyThrGlyThrGlyThrVal-----As	832
1419	QY	GCTAGAACACGAGCTTAGTAACTACATCGAATTGTAAACAGAAGATCTCCAAGCCCTGA	1478
832	Db	pIleLysAlaThrLeuGlyAsn-----Valise	841

QY 1479 AAATCTGTCTAGTAACTCGGAGAAATCCCTAACTCTTATCTGAAGTAGTCTTA 1533

Db 841 rSerGlnValSerArgLeuSerValThrAlaAlaGluLeuLeuValLeu 859

RESULT 8

```

US-11-033-039-1244
; Sequence 1244, Application US/11033039
; Publication No. US2006002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245, 871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197, 000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396, 813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1244
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Influenza A virus
US-11-033-039-1244

```

Alignment Scores:	
Pred. No.:	0.165
Pred. Score:	100.00
Percent Similarity:	31.0%
Best Local Similarity:	20.1%
Query Match:	2.9%
DB:	7
Length:	566
Matches:	118
Conservative:	64
Mismatches:	165
Indels:	240
Gaps:	33

US-10-723-552-3 COPY 5620 7533 (1-1914) X US-11-033-039-1244 (1-566)

[illegible]


```
Db 603 PheAspIleThrAsnTrpLeuTrpTyrIleArgLeuPheIleMetIleValGlyGlyLeu 622
Qy 1774 ATAGTCCTCCCTCTACTACAGTTCGGGCATGATTATTAAACAAGTTAATTGCCTTC 1833
Db 623 IleGlyLeuArgIleValPheAlaVal---LeuSerIleValAsnArg----- 637
Qy 1834 ATTAGAACGAATAAGTCGATCCAGATCATGTGTACTTAGACAACAGTACCAAGCCCG 1893
Db 638 ValArgGlnGlyTyrSerProLeuSerPheGlnThrLeuLeuProAlaSerArgGlyPro 657
Qy 1894 TCTAGCAGGGAA 1905
Db 658 AspArgProGlu 661

RESULT 12
US-11-014-842A-27
; Sequence 27, Application US/11014842A
; Publication No. US20060013826A1
; GENERAL INFORMATION:
; APPLICANT: TANGY, FREDERIC
; APPLICANT: LORIN, CLARISSE
; APPLICANT: MOLLET, LUCILE
; APPLICANT: DELEBEQUE, FREDERIC
; TITLE OF INVENTION: RECOMBINANT MEASLES VIRUSES EXPRESSING EPITOPES OF
; TITLE OF INVENTION: ANTIGENS OF RNA VIRUSES - USE FOR THE PREPARATION OF
; TITLE OF INVENTION: VACCINE COMPOSITIONS
; FILE REFERENCE: 2356.0093
; CURRENT APPLICATION NUMBER: US/11/014.842A
; CURRENT FILING DATE: 2004-12-20
; PRIOR APPLICATION NUMBER: PCT/EP03/07146
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: EP 02291550.8
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn ver. 3.3
; SEQ ID NO 27
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-014-842A-27

Alignment Scores:
Pred. No.: 0.399 Length: 869
Score: 96.50 Matches: 100
Percent Similarity: 30.9% Conservative: 62
Best Local Similarity: 19.1% Mismatches: 185
Query Match: 2.8% Indels: 177
DB: 7 Gaps: 22

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-11-014-842A-27 (1-869)
Qy 769 GGACGAGACCATCTCTAACAATA-----ACTTCGTGATCAGCCCACT 813
Db 238 GlySerGlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProVal 257
Qy 814 GAGTCTAACACGACGACTAAATCGGGCA----- 843
Db 258 ValSerThrGlnLeuLeuLeuAsnGlySerLeuAlaGluLeuAspIleValIleArgSer 277
Qy 844 -----AAACTTTTGTAGCCTCATCCAGGAGCTTTTCAAGCTCTT 882
Db 278 GluAspPheThrAspAsnValLysThrIleIleValGlnLeuAsnGluSerValValIle 297
Qy 883 AACTCCAGACTCCAGAGGCTCTCTCTCTTTGTTGGCTATGCTTAGCTTCGGGCCCACT 942
Db 298 AsnCysThrArgProAsnAsnThrArgGluArgLeuSerIleGlyProGlyArgAla 317
Qy 943 TACTATGAA-----GGAATGGCTAGAGAGGAATTCATGTGCACAAA 987
Db 318 PheTyrAlaArgArgAsnIleIleGlyAspIleArgGlnAlaHisCysAsnIleSerArg 337
Qy 988 GAACATAGAGACCAATGCATGGGATCCCAAAATAAGCTTACCCTTACTGAGTTTCT 1047
```

```
Db 338 -----AlaLysTrpAsnAsn-----ThrLeuGlnGlnIle--- 347
Qy 1048 GGAAGGCGACCTCATAGGAAGGTTCCCCATCCCAACACACCTTTTGTAAACACACT 1107
Db 348 -----ValIleLysLeuArgGluLysPheArgAsnLysThr 359
Qy 1108 GAAGCCTTTTAACTCAAACTCT-----GAGAGTCAATATCTGGTACCTGGTTATGACAGG 1161
Db 360 IleAlaPheAsnGlnSerSerGlyAspProGluIleValMetHisSerPheAsn--- 378
Qy 1162 TGTGGGCATGTAATACTGGATTAAACCCCTTGTGTTCCACCTGGTTTAAACCAAACT 1221
Db 379 -----CysGlyGlyGluPhePheTyrCysAsnThrAlaGlnLeuPheAsnSerThr 395
Qy 1221 ----- 1221
Db 396 TrpAsnValAlaGlyThrAsnGlyThrGluGlyAsnAspIleIleThrLeuGlnCys 415
Qy 1222 -----AAAGATTTTTCATTTATGGTCCAAATTTGTCCTCCAGAGTGTATTACTATCCC--- 1272
Db 416 ArgIleLysGlnIleIleAsnMetTrpGlnLysValGlyLysAlaMetTyrAlaProPro 435
Qy 1272 ----- 1272
Db 436 IleThrGlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGly 455
Qy 1273 -----GAAAAGCAATCCTT-----GATGAATATGAC 1299
Db 456 GlyAsnSerThrGluThrGluThrGluIlePheArgProGlyGlyGlyAspMetArgAsp 475
Qy 1300 -----TACAGAAATCATCGACAAAAGAGAGAACCATATCTCTGACA 1341
Db 476 AsnTrpArgSerGluLeuTyrLysValValArgIleGluProIleGlyValAla 495
Qy 1342 CTT-----GCT 1347
Db 496 ProThrArgAlaLysArgArgThrValGlnArgGluLysArgAlaValGlyIleGlyAla 515
Qy 1348 GTGATGCTCGGACTTGGAGTGGCGACAGGTGTAGGAACAGACAGCTGCC----- 1398
Db 516 ValPheLeuGlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaAlaSerValThrLeu 535
Qy 1399 -----CTGGTCACGGGACACAGCAGCTAGAAAACAGGACTT----- 1434
Db 536 ThrValGlnAlaArgLeuLeuLeuSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArg 555
Qy 1435 -----AGTAACCTACATCGAATTGTAAACAGAGAATCTCCAAAGCCCTAGAA 1479
Db 556 AlaIleGluAlaGlnGlnAsnMetLeuArgLeuThrValTrpGlyIleLysGlnLeuGln 575
Qy 1480 AAATCTGTCTAGTAACCTGGAGGAATCCCTA----- 1509
Db 576 AlaArgValLeuAlaLeuGluArgTyrLeuArgAspGlnGlnLeuMetGlyIleTrpGly 595
Qy 1510 -----ACCTCTCTATCTGAAGTAGCTTACAGAATAGAGA 1545
Db 596 CysSerGlyLysLeuIleCysThrThrSerValProTrpAsnValSerTrpSerAsnLys 615
Qy 1546 GGGTTAGATTATTATTCTTAAAGAGGAGGATTATGTGTAGCTTTGAAGAGGAGATGTC 1605
Db 616 SerValAspAspIleTrp-----AsnAsnMetThrTrpMetGluTrpGluArgGluIle 633
Qy 1606 TGTTTTTATGTGATCATTCAGGGCCATCAGAGACTCCATGAAACAGCTTAGAGAAAGG 1665
Db 634 AspAsnTyrThrAspTyr-----IleTyrAspLeuLeuGluLysSerGlnThrGln 650
Qy 1666 TTCGAGAAGCGTCGAGGGGAAAAGAAACTACTCAAGGGTGGTTTGAG-----GGATGG 1719
Db 651 GlnGlnLysAsnGlnLysGluLeuLeuGluLeuAspLysTrpAlaSerLeuTrpAsnTrp 670
Qy 1720 TTCAACAGAGTCT-----CTTTGGTTGGCTACCTTCTCTGCTTTAAACAGGACCTTTA 1773
```

Db 671 PheAspIleThrAsnTrpLeuTrpTyrlleArgLeuPheIleMetIleValIleGlyLeu 690
Qy 1774 ATAGTCCTCCTCTGTTACTCACAGTTCGGCCATGATTATTAAACAAGTTAATTGCCTTC 1833
Db 691 IleGlyLeuArgIleValPheAlaVal---LeuSerIleValAsnArg----- 705
Qy 1834 ATTAGACGAATAAGTCAGTCACAGTCCAGATCATGCTACTTAGACACACATCAAGCCCG 1893
Db 706 ValArgGlnGlyTyrSerProLeuSerPheGlnThrLeuLeuProAlaSerArgGlyPro 725
Qy 1894 TCTAGCAGGAA 1905
Db 726 AspArgProGlu 729

RESULT 13

US-11-175-690-528
; Sequence 528, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 528
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-528

Alignment Scores:

Pred. No.:	0.424	Length:	742
Score:	96.00	Matches:	64
Percent Similarity:	32.0%	Conservative:	34
Best Local Similarity:	20.9%	Mismatches:	88
Query Match:	2.8%	Indels:	120
DB:	7	Gaps:	14

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-11-175-690-528 (1-742)

Qy 1030 ACCCTTACTAGGTTTCTGAAAGGCACCTGCATGAGAAAGGTTTCCCCATCCCAACAA 1089
Db 446 ThrLeuValGluValSerArgAsn-----LeuGlyLysVal-----GlySer 459
Qy 1090 CACCTTTGTACCACTGAGCCCTTTAATCAACCTCTGAGAGTCATATCTG----- 1143
Db 460 LysCysCysLysHisProGluAlaLysArgMetProCysAlaGluAspTyrLeuSerVal 479
Qy 1144 -----GTACTGGTTATGACAGTGGTGGGCA 1170
Db 480 ValLeuAsnGlnLeuCysValLeuHisGluLysThrProValSerAspArgValThrLys 499
Qy 1171 TGTAATACT-----GGATTAAACCCCTGTGTTCACCTCTGGTGTTCACCA 1218
Db 500 CysCysThrGluSerLeuValAsnArgProCysPheSerAlaLeuGluValAspGlu 519

Qy 1219 ACT-----AAAAGATTTTGCATT 1236
Db 520 ThrTyrValProLysGluPheAsnAlaGluThrPheThrPheHisAlaAspIleCysThr 539
Qy 1237 ATGTCCTCAA-----ATTGTTCCCGAGTGTATTATC 1266
Db 540 LeuSerGluLysGluArgGlnIleLysLysGlnThrAlaLeuValGluLeuValLysHis 559
Qy 1267 TATCCCGAA-----AAAGCATCTCTGTGATAT- 1296
Db 560 LysProLysAlaThrLysGluGlnLeuLysAlaValMetAspAspPheAlaAlaPheVal 579
Qy 1297 -----GACTACAGAAATCATCGACAAAAGAGACCACTATCTCTG 1338
Db 580 GluLysCysCysLysAlaAspAspLysGluThrCysPheAlaGluGluGlyLysLysLeu 599
Qy 1339 ACACCTGCTGTGATGCTCGGACTTGGAGTGGCAGCAGGTGTAGGAACAGCACTGCC 1398
Db 600 ValAlaAlaSerGlnAlaLeuGlyLeuAlaPro-----ThrSerSerSerThr 616
Qy 1399 -----CTGCTACGGGACCACAGCAGCTAGAAACAGGA 1431
Db 617 LysLysThrGlnLeuGlnLeuGluHisLeuLeuLeuAspLeuGlnMetIleLeuAsnGly 636
Qy 1432 CTTAGTAAC-----CTACATCGAATTGTAAACA----- 1458
Db 637 IleAsnAsnTyrLysAsnProLysLeuThrArgMetLeuThrPheLysPheTyrMetPro 656
Qy 1459 -----GAAGATCTCAAGCCCTAGAAAAATCTGTCAGTAACTGGAG 1500
Db 657 LysLysAlaThrGluLeuLysHisLeuGlnCysLeuGluGluGluLysProLeuGlu 676
Qy 1501 GAA-----TCCTTAACCTCC 1515
Db 677 GluValLeuAsnLeuAlaGlnSerLysAsnPheHisLeuArgProArgAspLeuIleSer 696
Qy 1516 TTATCTGAAGTAGTCCTACAGAAATAGAGAGGGTTAGATTTATTATTCTTAAAAAGAGGA 1575
Db 697 AsnIleAsnValIleValLeuGluLeuLysGlySerGluThrThrPheMet----- 713
Qy 1576 GGATTATGTGTAGCCTTGAAGGAGGAATGCTGTTTTTATGTGTGATCATTCAGGGGCCATC 1635
Db 714 -----CysGluTyrAlaAspGluThrAlaThrIle 723
Qy 1636 AGAGACTCCATGAACAAG 1653
Db 724 ValGluPheLeuAsnArg 729
RESULT 14
US-11-175-690-534
; Sequence 534, Application US/11175690
; Publication No. US20060014254A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF605
; CURRENT APPLICATION NUMBER: US/11/175,690
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: PCT/US04/001369
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 60/441,305
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/453,201
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US 60/467,222
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/472,816
; PRIOR FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: US 60/476,267
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/505,172
; PRIOR FILING DATE: 2003-09-24

```
; PRIOR APPLICATION NUMBER: US 60/506,746
; PRIOR FILING DATE: 2003-09-30
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 534
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-175-690-534

Alignment Scores:
Pred. No.: 0.424 Length: 742
Score: 96.00 Matches: 64
Percent Similarity: 32.0% Conservative: 34
Best Local Similarity: 20.9% Mismatches: 88
Query Match: 2.8% Indels: 120
DB: Gaps: 14

US-10-723-552-3_COPY_5620_7533 (1-1914) x US-11-175-690-534 (1-742)
QY 1030 ACCCTTACTGAGTTTCTGAAAGGCACCTGCATAGAAAGGTTCCCCATCCACCAA 1089
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 446 ThrLeuValGluValSerArgAsn-----LeuGlyLysVal-----GlySer 459

QY 1090 CACCTTTGTAAACACACATGAAGCCTTAAATCAACCTCTGAGAGTCAATATCTG----- 1143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 460 LysCysCysLysHisProGluAlaLysArgMetProCysAlaGluAspTyrLeuSerVal 479

QY 1144 -----GTACCTGGTTATGACAGTGTGGGCA 1170
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 ValLeuAsnGlnLeuCysValLeuHisGluLysThrProValSerAspArgValThrLys 499

QY 1171 TGTAAATACT-----GGATTAACCCCTTGTGTTCCACCTTGGTTTAAACCAA 1218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 CysCysThrGluSerLeuValAsnArgArgProCysPheSerAlaLeuValAspGlu 519

QY 1219 ACT-----AAAGATTTTGCATT 1236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 520 ThrTyrValProLysGluPheAsnAlaGluThrPheHisAlaAspIleCysThr 539

QY 1237 ATGCTCAA-----ATTGTTCCCGAGTGTATTAC 1266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 LeuSerGluLysGluArgGlnIleLysLysGlnThrAlaLeuValGluValLysHis 559

QY 1267 TATCCCGAA-----AAAGCAATCCTTGATGAATAT----- 1296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 LysProLysAlaThrLysGluLysAlaValMetAspAspPheAlaAlaPheVal 579

QY 1297 -----GACTACAGAAATCATCGAAAAAGAGAGAACCCATATCTCTG 1338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 580 GluLysCysCysLysAlaAspAspLysGluThrCysPheAlaGluGluGlyLysLysLeu 599

QY 1339 ACATCTGCTGATGCTCGGACTTGGAGTGCACAGGTGTAGAAACAGGACAGCTGCC 1398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 600 ValAlaAlaSerGlnAlaLeuGlyLeuAlaPro-----ThrSerSerSerThr 616

QY 1399 -----CTGTCTACCGGACCACACAGCAGCTAGAAACAGGA 1431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 617 LysLysThrGlnLeuGlnLeuGluHisLeuLeuAspLeuGlnMetIleLeuAsnGly 636

QY 1432 CTTAGTAAC-----CTACATCGAATTGTAAACA----- 1458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 637 IleAsnAsnTyrLysAsnProLysLeuThrArgMetLeuThrPheLysPheTyrMetPro 656

QY 1459 -----GAAGATCTCGAGCCCTAGAAAAATCTGTCAGTAACCTGGAG 1500
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 657 LysLysAlaThrGluLeuLysHisLeuGlnCysLeuGluGluLysProLeuGlu 676

QY 1501 GAA-----TCCCTAACCTCC 1515
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 677 GluValLeuAsnLeuAlaGlnSerLysAsnPheHisLeuArgProArgAspLeuIleSer 696

QY 1516 TTATCTGAAGTAGTCTTACAGAATAGAGAGGGTTAGATTATTATTCTTAAAGAGGA 1575
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

